



## 저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

공학석사 학위논문

심층 오토인코더를 결합한 신경망 기법의  
채널저류층 특성화 연구

**Characterization of Channelized Reservoir Using  
the Neural Network Incorporated with Deep Autoencoder**

2017 년 2 월

서울대학교 대학원

에너지시스템공학부

안 성 인

## 초 록

이 연구는 심층 오토인코더를 결합한 신경망 기법을 활용하여 유체투과도와 생산이력을 통합 분석함으로써 불균질 저류층의 미래 생산량을 신뢰도 높게 예측하였다. 기존 신경망 기반의 히스토리 매칭은 정적 및 동적자료의 비선형관계를 단순화하였기 때문에 생산이력의 매칭효율이 높다 하더라도 물성의 불균질한 분포 특성을 보존하기 어려웠다. 이 불균질 특성화의 실패로 인해 히스토리 매칭 이후에도 생산추이 예측의 일관성을 확보하지 못하였다. 이 연구는 심층 오토인코더를 이용한 신경망 체계를 개발하여 생산정의 생산이력(생산정별 물, 오일 생산추이)과 유체투과도 분포의 비선형적 상관관계를 학습하였다. 대상 저류층은 3개의 생산정과 1개의 물 주입정이 존재하는 불균질 유체투과도의 채널저류층이다. 개발한 기법의 성능은 대상 저류층을 포함한 경우와 그렇지 않은 경우의 학습에 대해 일반 신경망기법과 비교하여 생산이력과 유체투과도 분포의 예측정확도 및 연산시간을 평가하였다. 대상 저류층을 포함한 경우 두 기법 모두 사암의 유체투과도와 채널연결성은 근접하게 예측하였지만, 일반 신경망 기법의 경우 세일성 사암의 유체투과도 분포 예측에 실패하여 생산이력 오차가 32%로 높게 나타났다. 반면, 개발한 신경망 기법은 세일성 사암의 유체투과도 또한 양호하게 예측하여 생산이력 오차가 8.6%로 크게 감소하였다. 또한 대상 저류층이 포함되지 않은 경우에도 개발한 신경망 기법은 생산이력 오차가 15.6%로 일반 신경망기법보다 연산시간이 약 14배 정도 빠름에도 불구하고 오차는 1/3수준으로 매우 양호하였다. 불균질 분포 예측에서 일반 신경망기법은 유체투과도 분포와 사암채널의 분포특성과의 상관성을 규명하지 못한 반면, 개발한 기법은 유체투과도와 채널의 분포 모두 대상저류층과 유사하였다. 개발한 심층 오토인코더 기반의 신경망 기법은 스케일이 다양한 생산정 별 생산이력과 저류층 물성의 불균질도를 통합하여 신뢰도 높은 생산추이 추정이 가능하여 생산설계 과정의 의사결정을 위한 효용성 높은 도구로 활용될 수 있다.

**주요어:** 신경망 기법, 인공신경망, 심층 오토인코더, 채널저류층

**학번:** 2015-21320

# 목 차

초 록.....	i
List of Tables .....	iv
List of Figures.....	v
1. 서론.....	1
2. 이론적 배경 .....	6
2.1. 인공신경망(Artificial Neural Network) .....	6
2.2. 심층신경망(Deep Neural Network).....	10
2.3. 다점지구통계학(Multiple-point Geostatistics).....	13
3. 연구 방법 .....	15
3.1. 신경망을 이용한 앙상블 학습 .....	15
3.2. 심층 오토인코더의 적용.....	21
4. 연구 결과.....	24
4.1. 참조모델 구축과 신경망 기법 검증 .....	25
4.2. 신경망 기법의 예측성능 평가 .....	43

4.3. 균질한 유체투과도 모델과의 비교.....	54
<b>5. 결론.....</b>	<b>59</b>
<b>참고문헌.....</b>	<b>61</b>
<b>Abstract.....</b>	<b>65</b>
부록 A. 심층 오토인코더의 은닉층 뉴런수의 결정 .....	67
부록 B. 신경망 기법의 은닉층 뉴런수에 따른 성능.....	69

## **List of Tables**

<b>Table 4.1</b>	<b>Parameters for Training image .....</b>	<b>26</b>
<b>Table 4.2</b>	<b>Parameters for SNESim simulation.....</b>	<b>27</b>
<b>Table 4.3</b>	<b>Well information and boundary conditions of the reference field and initial ensembles .....</b>	<b>28</b>
<b>Table 4.4</b>	<b>Petrophysical parameters for reference field and initial ensembles .....</b>	<b>29</b>
<b>Table 4.5</b>	<b>Neuron descriptions for the application of artificial neural network ....</b>	<b>32</b>
<b>Table 4.6</b>	<b>Neuron descriptions for the application of propsed neural network using autoencoder .....</b>	<b>38</b>
<b>Table 4.7</b>	<b>Comparison of average computation time .....</b>	<b>57</b>

## List of Figures

<b>Figure 2.1</b>	<b>Comparison between biological neuron and artificial neuron: (a) human neuron (b) ANN neuron or hidden unity (c) human synapse (d) ANN synapses (Matarollo et al., 2013).....</b>	<b>6</b>
<b>Figure 2.2</b>	<b>Activation functions used in ANN (Hagan et al., 1996). ....</b>	<b>8</b>
<b>Figure 2.3</b>	<b>The algorithmic step of the BP algorithm (a) feed forward simulation (b) error estimation (c) back propagation (Araghinejad, 2014). ....</b>	<b>9</b>
<b>Figure 2.4</b>	<b>Example of compressed numerical coding of the inputs by autoencoder (Mitchell, 1997). ....</b>	<b>12</b>
<b>Figure 2.5</b>	<b>Training image and conditional probability in multiple-point geostatistics (최종근, 2013). . ....</b>	<b>14</b>
<b>Figure 3.1</b>	<b>Flow chart of the proposed method.....</b>	<b>16</b>
<b>Figure 3.2</b>	<b>Timing of training stop in data training and validation (Araghinejad, 2014). ....</b>	<b>19</b>
<b>Figure 3.3</b>	<b>Structure of artificial neural network.....</b>	<b>20</b>
<b>Figure 3.4</b>	<b>Application of deep autoencoder for transforming 2,500 grid permeability data to 100 hidden value data. ...</b>	<b>22</b>
<b>Figure 3.5</b>	<b>Proposed neural network structure with deep autoencoder. ....</b>	<b>23</b>

<b>Figure 4.1</b>	<b>Training Image.....</b>	<b>27</b>
<b>Figure 4.2</b>	<b>Reference field: (a) realization of SNESim (b) log-permeability distribution.....</b>	<b>28</b>
<b>Figure 4.3</b>	<b>Well daily production of the reference field: (a) Oil rate (b) Water rate. .....</b>	<b>30</b>
<b>Figure 4.4</b>	<b>Ensembles for neural network training: (a) reference (b) initial ensembles. .....</b>	<b>32</b>
<b>Figure 4.5</b>	<b>Log permeability distribution: (a) reference (b) predicted ensemble(P50) by ANN in case of including the reference. ....</b>	<b>34</b>
<b>Figure 4.6</b>	<b>Averaged error of permeability predicted by ANN in case of including the reference.....</b>	<b>34</b>
<b>Figure 4.7</b>	<b>Histogram of reference and predicted ensemble(P50) by ANN in case of including the reference.....</b>	<b>35</b>
<b>Figure 4.8</b>	<b>Well production rate of predicted ensembles with reference using ANN: (a) oil rate of P1 (b) water rate of P1 (c) oil rate of P2 (d) water rate of P2 (e) oil rate of P3 (f) water rate of P3. ....</b>	<b>36</b>
<b>Figure 4.9</b>	<b>Log permeability distribution: (a) reference (b) predicted ensemble(P50) by proposed NN in case of including the reference. ....</b>	<b>40</b>
<b>Figure 4.10</b>	<b>Averaged error of permeability predicted by proposed NN in case of including the reference.....</b>	<b>40</b>



<b>Figure 4.11 Histogram of reference and predicted ensemble(P50) by proposed NN in case of including the reference.....</b>	<b>41</b>
<b>Figure 4.12 Well production rate of predicted ensembles with reference using proposed NN: (a) oil rate of P1 (b) water rate of P1 (c) oil rate of P2 (d) water rate of P2 (e) oil rate of P3 (f) water rate of P3. ....</b>	<b>42</b>
<b>Figure 4.13 Log permeability distribution of predicted ensembles: (a) reference (b) ANN (c) proposed NN in case of excluding the reference. ....</b>	<b>45</b>
<b>Figure 4.14 In case of excluding the reference, averaged error of permeability predicted by: (a) ANN (b) proposed NN.....</b>	<b>46</b>
<b>Figure 4.15 In case of excluding the reference, histogram of reference and predicted ensemble(P50) by: (a) ANN (b) proposed NN. ....</b>	<b>47</b>
<b>Figure 4.16 Well production rate of predicted ensembles without reference using ANN: (a) oil rate of P1 (b) water rate of P1 (c) oil rate of P2 (d) water rate of P2 (e) oil rate of P3 (f) water rate of P3.....</b>	<b>49</b>
<b>Figure 4.17 Well production rate of predicted ensembles without reference using proposed NN: (a) oil rate of P1 (b) water rate of P1 (c) oil rate of P2 (d) water rate of P2 (e) oil rate of P3 (f) water rate of P3. ....</b>	<b>50</b>
<b>Figure 4.18 Comparison to boxplot of daily well oil productions error between ANN and proposed NN in case of including and excluding the reference. ..</b>	<b>53</b>
<b>Figure 4.19 Comparison to boxplot of uncertainty range of cumulative oil productions</b>	

at 1,000 days between ANN and proposed NN in case of including and excluding the reference.....	53
Figure 4.20 Log permeability distribution: (a) reference (b) predicted ensemble(P50) from a homogenous model.....	54
Figure 4.21 Well production rate of predicted ensembles from homogeneous model: (a) oil rate of P1 (b) water rate of P1 (c) oil rate of P2 (d) water rate of P2 (e) oil rate of P3 (f) water rate of P3. ....	55
Figure 4.22 Comparison of boxplot of daily well oil productions error.....	58
Figure 4.23 Comparison of boxplot of uncertainty range of cumulative oil productions at 1,000 days. ....	58
Figure A.1 Error variance and computing time depending on the number of hidden neurons in deep autoencoder.....	68
Figure B.1 Comparison of correlation coefficient and computing time between ANN and proposed NN. ....	70

## 1. 서론

석유 개발을 위한 E&P사업은 다른 사업과 비교하여 사업에 필요한 비용이 크고 사업 기간이 길기 때문에 잘못된 의사결정은 막대한 손해를 초래할 수 있다. 해외 자원개발에 따른 정치적 불확실성, 사업에 필요한 자본 투자, 유치, 회수에 따른 경제적 불확실성, 지하에 존재하는 석유자원의 특성으로 인하여 저류층과 유체유동 분석 등이 어려운 기술적 불확실성은 의사결정 과정에서 어려움을 야기한다. 그러므로 성공적인 E&P사업을 위해서는 주어진 자료를 활용하여 시황에 따른 탐사 및 개발 사업의 시작 여부, 시추계획, 생산 계획, 지분매각 등 사업성속도에 따른 수많은 의사결정을 올바르게 해야 한다. 만약 신뢰할 만한 저류층 모델을 구축하여 미래 생산량을 예측할 수 있다면 이 또한 합리적인 의사결정이라고 볼 수 있다. 저류층 모델의 구축을 위해서는 저류층 특성값에 대한 규명이 필요하며 이를 위해 저류층의 특성화 과정이 수행되어야 한다.

저류층 특성화는 정적자료와 동적자료를 이용하여 수행할 수 있다. 정적자료는 시간에 따라 변하지 않는 자료로 코어시료의 정보, 로깅자료 등이 있다. 정적자료를 이용하여 초기 저류층 모델을 구축할 수 있지만 주어진 정보는 한정되어 있기 때문에 저류층의 상태를 신뢰할 수 있게 규명하지 못한다. 따라서 생산정의 생산자료나 압력자료와 같이 시간에 따라 변화하는 동적자료를 결합한 저류층 특성화가 수행되어야 한다. 초기 저류층 모델에 동적자료를 결합하는 과정은 관측기간 동안의 생산이력과 저류층 모델의 생산량 오차를 최소화 하는 과정으로 히스토리 매칭이라고도 한다.

히스토리 매칭의 중요한 목적은 신뢰할 수 있는 미래생산추이를 예측하기 위해서 관측된 동적 자료를 이용한 특성화를 통해 저류층 시물레이션의 성능을 향상 시키는 것이다. 고전적인 히스토리 매칭에서는 직접 인자를 시행착오를 거쳐 조정하면서 반복적이고 많은 시간이 소모되었다. 반면 최근에는 정적자료와 동적자료의 비선형적인 관계를 예측하고 적은 자원으로 빠르게 히스토리 매칭을 수행하기 위한 연구가 진행되고 있었으며, 인공신경망(ANN; Artificial neural network)은 입력자료와 출력자료가 비선형관계를 가지는 문제에 적용할 수 있는 도구(Zupan, 1994)로서 히스토리 매칭에 적용한 연구가 활발하게 진행되었다. Cullick 등(2006)은 165개의 생산정들이 백만 개의 격자에 포함된 모델에 대해 시물레이션을 수행하였을 때 소요되는 시간을 문제점으로 지적하면서 신경망 네트워크를 사용하여 히스토리 매칭에서 발생하는 비선형문제를 고려하면서도 빠른 연산이 가능한 프록시 네트워크를 연구하였다. 이 밖에도 히스토리 매칭의 불확실성을 정량화하기 위한 인자 설정과정에서 회귀모델과 신경망 네트워크를 비교 하는 등 히스토리 매칭의 성능을 향상시키기 위해 비선형 신경망 프록시 모델에 대한 연구가 활발히 진행되었다(Lechner and Zangl, 2005; Reis, 2006; Silva et al., 2007; Sampaio et al., 2009).

Ramgulam (2006)은 생산량 차이를 이용하여 저류층의 정적 물성(유체투과도, 공극률, 두께)을 예측하는 역산방법을 통해 히스토리 매칭을 위한 시물레이션 횟수를 감소시켰으며, 인공신경망의 은닉층 뉴런 숫자의 민감도 분석을 통해 예측성능을 향상 시켰다. 하지만 저류층 모델의 채널 연결성, 균열 등의 구조에 대한 정보를 사전에 가정하여 구역별 균질한(Homogeneous) 유체투과도를 가진다는 가정을 하여 문제를 지나치게 단순화 시켰기 때문에

불확실성이 큰 저류층에는 적용하지 못하는 한계점이 있다.

Shahkarami 등(2014)은 생산정의 생산량에 영향을 미치는 중요인자들에 대한 분석을 통해 각 인자들의 순위와 영향을 인공신경망을 통해 학습하여 시간과 공간적 자료의 통합 관리가 가능한 저류층 모델을 제시하였지만, 생산정 위치에서의 유체투과도 값을 예측하고 생산정의 영향반경을 가정하여 다른 격자의 유체투과도 값을 예측하였기 때문에 저류층 모델의 유체투과도 매칭이 생산정의 숫자에 영향을 받는 한계점이 있다.

또한 기계 학습의 분야로서 인공신경망과 함께 심층신경망(DNN; Deep neural network)이 석유공학에 적용되었다. 기존의 인공신경망과 달리 심층신경망은 다중 층 구조를 가지고 좀 더 복잡한 비선형관계를 예측할 수 있어 이미지 처리, 패턴인식, 회귀문제 등 많은 양의 자료로 구성 된 복잡한 모델을 예측하는 분야에 적용되고 있다. 저류층 특성화 연구와 관련하여 Korjani 등(2016)은 퍼지 크리깅(Fuzzy kriging)과 심층신경망 기법을 이용하여 저류층 모델의 검층기록(Well log) 예측 및 분석을 통해 층적 퇴적층의 암상분포를 예측하였다.

한편 채널저류층은 전통 퇴적 저류층 중의 한 형태로 미국, 캐나다, 베네수엘라 등 전 세계적으로 주요한 탄화수소 저류층으로 볼 수 있다. 이러한 채널 저류층은 사암과 셰일성 사암의 암상 분포에 따라 불균질한 횡적·종적 연결성을 가진다. 또한 채널 저류층에서의 유체흐름은 사암연결성에 따라 큰 영향을 받기 때문에, 유·가스전 개발에 있어 미래의 올바른 생산거동 예측을 통한 합리적인 의사 결정을 위해서는 이러한 사암연결성을 고려한 채널 저류층의 특성화가 필수적이다. 채널저류층의

특징을 고려한 특성화를 위해 많은 연구들은 저류층 모델의 격자 별 정보를 예측하면서도 객체 즉, 채널의 연결성을 보존할 수 있는 트레이닝 이미지(training image, TI)를 활용하였다(Daly, 2005; Jafarpour and Mclaughlin, 2007; Tavakoli et al., 2013; 이경복, 2014).

Jafarpour와 Mclaughlin (2009)은 두 가지의 TI로부터 초기 저류층 모델을 생성하여 다양한 채널폭과 방향을 가진 300개 이상의 모델을 생성하였다. 이러한 채널 모델에 EnKF(Ensemble Kalman Filter)를 적용하여 안정적인 히스토리 매칭을 수행하였다. EnKF는 다수의 저류층모델인 앙상블을 이용해 이들의 추정오차공분산을 최소화하면 해를 찾는 기법이다. 다수의 해집합을 이용하여 불확실성을 평가할 수 있을 뿐 아니라 확실한 수학적 기반을 가지고 있기 때문에 해양학, 기상학, 수리지질학 등에도 활용되며, 저류층 특성화 분야에서도 활발하게 연구하고 있다. 하지만 등가의 확률로 생성된 모델들이 충분히 많아야 하며, 저류층 물성값이 정규분포를 따라야 한다는 가정이 만족하여야 하는 한계점이 있다.

임서진 (2016)은 트레이닝 이미지의 불확실성을 고려하기 위해 채널 너비와 방향성이 다른 다수의 TI로부터 초기 앙상블을 생성하여 관측된 오일생산량과 예측한 생산량의 차이가 적은 TI를 선별하여 이들로부터 초기저류층 모델을 재생성함으로써 EnKF의 성능을 개선시켰다. 하지만 저류층 모델이 복잡할 경우 TI 선별과정에서 어려움이 있으며, 계산시간이 기존의 방법에 비하여 증가하는 한계점이 있다.

기존 연구에서 드러난 문제점들은 다음과 같다.

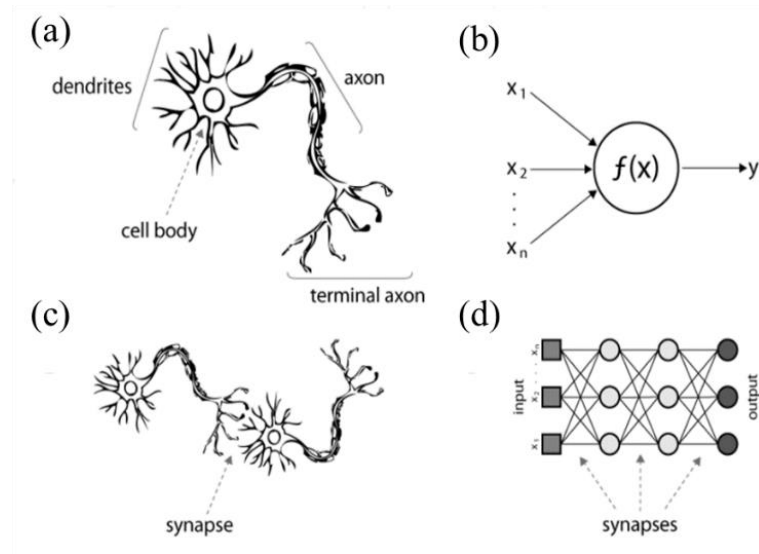
- 1) 저류층 특성화를 위해 기존 인공신경망을 적용하여 저류층의 물성값을 예측한 연구들의 경우, 저류층 모델의 구역별 물성(유체투과도, 공극률 등)의 평균값이나 생산정이 위치한 격자의 물성값만을 예측하는 문제로 단순화한 한계점이 있다.
- 2) 인공신경망을 이용하여 비선형성을 지닌 모델에서 수 개에서 수십 개가 아닌 수천 개 이상의 많은 자료를 처리해야 하는 경우 연산 시간이 크게 증가하고 성능이 감소하는 한계점이 있다.
- 3) TI와 EnKF를 이용하여 채널저류층을 특성화할 경우 히스토리 매칭에 초점을 맞추어 생산 이력은 잘 따르지만 유체투과도의 분포를 제대로 예측하지 못하는 현상이 발생하였고, 복잡한 참조모델의 경우 채널 연결성을 보존하지 못하는 한계점이 있다.

본 연구의 핵심 목적은 생산정의 위치뿐만 아니라 저류층 모델 전체의 비균질한 유체투과도 분포와 채널 연결성에 대한 불확실성을 평가할 수 있는 채널 저류층 특성화 기법의 제안이다. 이를 위해 트레이닝 이미지를 활용하여 채널 분포의 지질학적 불확실성을 고려하면서도 많은 수의 자료를 처리할 수 있는 심층 오토인코더가 적용된 저류층 특성화 기법을 개발하였다.

## 2. 이론적 배경

### 2.1. 인공신경망(Artificial Neural Network)

인공신경망은 생물학적 신경세포를 기반으로 수학적으로 구현한 모델로서 모든 인공신경망 방법에서는 신경세포(Neurons or hidden units)의 개념을 신경망 구조의 기초로 설정하며, 각각의 신경세포들은 생물학 영역의 시냅스(synapse)를 나타낸다. Figure 2.1은 생물학적 신경세포와 인공신경망의 뉴런을 비교한 것이다.



**Figure 2.1 Comparison between biological neuron and artificial neuron: (a) human neuron (b) ANN neuron or hidden unity (c) human synapse (d) ANN synapses (Maltarollo et al., 2013).**



각각의 은닉 요소(Hidden unity)들은 다음 세포로의 신경신호 전달을 조절하는 활성화함수로 구성되어 있으며, 또한 회귀식으로 구성되어 있기 때문에 비선형적 관계를 가지는 입력정보와 출력정보의 정보처리가 가능하다. 그러므로 만약 하나 이상의 신경세포로 구성 된 인공신경망을 사용한다면 비선형적인 상관관계를 가지는 문제를 해결할 수 있으며, 이러한 특징으로 인해 한편에서는 인공신경망을 ‘Black box’의 개념과 같다고 표현한다 (Maltarollo et al., 2013).

인공신경망은 입력층(Input layer), 출력층(Output layer), 은닉층(Hidden layer)으로 최소 3개 이상의 층으로 구성되어 있으며, 각 층의 뉴런에서는 가중치의 합과 같은 단순한 계산만을 수행하고 다량의 뉴런들이 병행적으로 계산을 수행함으로써 연산을 수행할 있다. 두 뉴런 간의 연결 강도를 의미하는 가중치는 2차원 배열로 처리하여 입출력에 대한 각각의 가중치를 설정하며, 가중치의 합은 연산자를 통하여 활성화 함수를 사용하여 구현 된다. 인공신경망에서 사용하는 활성화 함수는 문턱형(Threshold), 선형(Linear), S자형(Sigmoid) 등이 있으며, Figure 2.2는 대표적인 활성화 함수를 나타낸 것이다. 또한 활성화 함수의 임계값 변동을 흡수하기 위한 가중치인 편향(Bias)이 있으며, 편향 값의 입력은 항상 1이며, 학습에 따라 가중치 값만 달라지게 된다.

인공신경망 기법은 신경망구조나 뉴런의 연결패턴에 따라 분류될 수 있다. 피드포워드 신경망(Feed forward networks)는 신경망 층 사이의 연결이 한쪽 방향으로만 이루어져 있어, 입력층에서부터 출력층까지 단방향으로 신호가 전달된다. 반면 피드백 신경망(Feedback networks) 또는 순환신경망

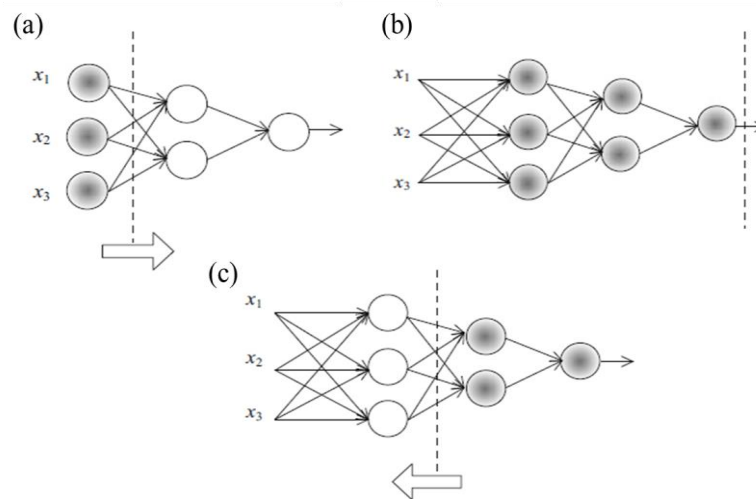
(Recurrent networks)기법은 신경망 층 사이의 연결이 양방향으로 이루어져 있기 때문에 피드백과정을 거쳐 전체적인 과정이 순환되는 형태를 가진다.

Name	Input/Output Relation	Icon	MATLAB Function
Hard Limit	$a = 0 \quad n < 0$ $a = 1 \quad n \geq 0$		hardlim
Symmetrical Hard Limit	$a = -1 \quad n < 0$ $a = +1 \quad n \geq 0$		hardlims
Linear	$a = n$		purelin
Saturating Linear	$a = 0 \quad n < 0$ $a = n \quad 0 \leq n \leq 1$ $a = 1 \quad n > 1$		satlin
Symmetric Saturating Linear	$a = -1 \quad n < -1$ $a = n \quad -1 \leq n \leq 1$ $a = 1 \quad n > 1$		satlins
Log-Sigmoid	$a = \frac{1}{1 + e^{-n}}$		logsig
Hyperbolic Tangent Sigmoid	$a = \frac{e^n - e^{-n}}{e^n + e^{-n}}$		tansig
Positive Linear	$a = 0 \quad n < 0$ $a = n \quad 0 \leq n$		poslin
Competitive	$a = 1 \quad \text{neuron with max } n$ $a = 0 \quad \text{all other neurons}$		compet

**Figure 2.2 Activation functions used in ANN (Hagan et al., 1996).**

인공신경망 기법에서 중요한 역할을 하는 단계는 학습단계이다. 학습단계는 예측 오차를 최소화하는 목표를 위해 가중치를 조절하는 과정이다. 학습 알고리즘에는 Conjugate gradient descent, quasi-Newton, Levenberg-Marquardt 등이 있으며, 대표적으로 사용되는 것은 역전파(Back

propagation)알고리즘이다. 역전파 알고리즘은 전향 단계와 후향 단계로 이루어지며 전향 단계에서 초기에 임의로 정한 연결 강도에 의해 신경망의 출력이 계산하고 후향 단계에서 연결 강도의 갱신이 이루어진다. 이 때문에 전향 단계에서의 출력 값은 목표 값과 오차를 발생시키게 되며, 후향 단계에서 목표 출력 값과 계산된 출력 값의 차이를 계산하여 오차를 구하고, 이를 역방향으로 연결 강도를 갱신하는 과정을 진행한다. 갱신된 연결강도를 기준으로 다시 알고리즘을 반복하면 더 작은 오차를 얻게 되고 이러한 오차가 원하는 값에 수렴할 때까지 반복 실행하게 된다. 연결강도의 보정과정에서 보정 값을 오차에 비례하게 하기 위해 일반화한 델타 규칙(Generalized delta rule)을 적용한다. Figure 2.3은 인공신경망 기법의 역전파 알고리즘의 흐름도를 나타낸 것이다.



**Figure 2.3 The algorithmic step of the BP algorithm (a) feed forward simulation (b) error estimation (c) back propagation (Araghinejad, 2014).**

이러한 인공신경망은 정보의 병렬분산처리가 가능하여 빠른 연산속도를 가지며, 연결강도를 조정하는 학습과정을 통해 정보를 추가하거나 변경할 수 있는 적응성을 가진다. 이에 따라 인공신경망은 경영, 금융, 의학, 안전, 에너지 분야 등 많은 영역에서 활발하게 적용되고 있으며, 석유공학에서도 다양한 분야에 적용되고 있다.

## 2.2. 심층신경망(Deep Neural Network)

보편적으로 딥러닝으로 명칭되는 심층 구조 학습은 2006년 이후 기계 학습 연구 분야의 새로운 영역으로 떠올랐다(Hinton et al., 2006). 심층신경망은 다양한 정의가 있지만 가장 보편적으로 패턴의 분석이나 분류, 특징의 추출이나 변형처리와 같은 비선형 문제를 지도학습(Supervised learning)이나 비지도학습(Unsupervised learning)을 통해 많은 수의 은닉층을 활용하여 정보를 처리하는 기계학습 기법으로 정의한다.

심층신경망은 층마다 다른 층위의 특징을 학습할 수 있으며, 낮은 층위의 단순하고 구체적인 특징에서부터 복잡하고 추상적인 더 높은 층위의 특징을 추출함으로써 데이터의 잠재적인 구조를 파악할 수 있다. 이러한 심층신경망이 최근 급격하게 이슈화 되고 연구될 수 있었던 이유는 이미지 처리 기술능력의 발달, 컴퓨터 하드웨어의 개발에 따른 연산비용 감소, 새로운 기계학습 정보처리 기법이 있다. 심층신경망이 활발하게 연구됨에 따라 복잡하고 비선형함수로 구성된 문제를 효과적으로 해결할 수 있고,

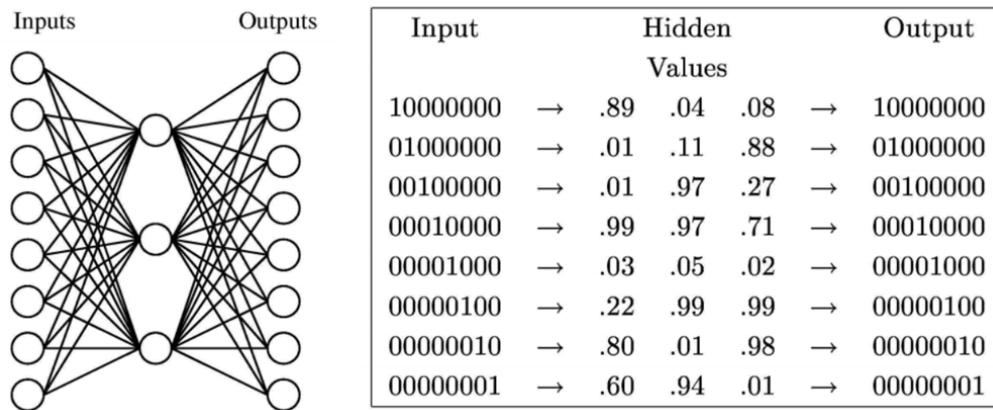
분류되지 않은 자료에 대한 학습성능도 뛰어나기 때문에 인공지능뿐만 아니라 그래픽 모델링, 최적화, 패턴인식, 신호처리 등 다양한 분야에서 활용되고 있다(Li and Yu, 2014).

### 2.2.1. 심층 오토인코더(Deep Autoencoder)

오토인코더(Autoencoder)는 심층신경망 기법 중 하나로 출력 벡터값들이 입력 벡터값들과 같은 차원을 가지는 특징이 있다. 주로 입력자료의 묘사를 위한 학습이나 입력 벡터값의 원본 자료를 은닉층에 코드화(encoding)하기 위해 사용된다.

오토인코더는 전형적으로 원본자료로 구성된 입력층과 변형된 특징을 나타내는 은닉층 그리고 입력층을 재구현하는 출력층으로 구성된다. 은닉층의 수가 하나 이상일 경우 심층 오토인코더라고 하며, 은닉층의 차원은 층의 뉴런 수를 통해 조절 가능하기 때문에 추출하고자 하는 특징을 목적에 따라 변형시킬 수 있다. 만약 자료의 특징을 압축하는 것이 목적이라면 은닉층의 뉴런 수를 적게 할 경우 낮은 차원의 값으로 변형시킬 수 있다(Li and Yu, 2014). Figure 2.4는 8자리의 2진수 데이터에 오토인코더를 적용한 예시이다. 8자리의 2진수를 나타내기 위해 입력층은 8개의 뉴런으로 구성되었으며, 출력층 또한 입력층을 재구현하기 위하여 8개의 뉴런으로 구성되었다. 이러한 8개의 자료값을 3개의 뉴런으로 구성된 은닉층을 통해 3개의 은닉값으로 변형되어 이진수 1,0,0,0,0,0,0,0의 8개의 값이 0.89, 0.04, 0.08의 값으로 코드화 되었다. 이와 같은 과정을 통해 입력층의 자료 값은

은닉층의 뉴런에 압축되어 코드화 될 수 있으며, 압축과 반대로 만약 높은 차원의 값으로 확대 변형하는 것이 목적이라면 은닉층의 뉴런 수를 늘림으로써 차원을 확대시킬 수 있다.

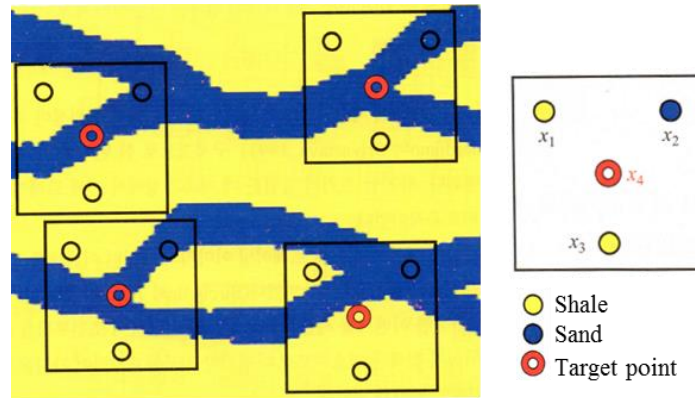


**Figure 2.4 Example of compressed numerical coding of the inputs by autoencoder (Mitchell, 1997).**

기존의 인공신경망 기법에서 많은 수의 은닉층에 역전파 알고리즘을 사용하는 데에는 큰 문제점들이 있다. 많은 수의 은닉층을 가진 구조에서 일단 오류가 시작층으로 역전파 될 경우 가중치의 변화는 미미하게 변할 뿐이고 이에 학습 성능 개선되지 못한다. 물론 이러한 문제를 해결하기 위하여 개선된 역전파 기법이 개발 되었지만 제한된 학습자료의 양에 따라 여전히 느리고 좋지 못한 성능을 보였다. 하지만 이러한 문제를 보완하기 위해 각 층의 뉴런을 사전 학습하는 오토인코더 기법이 제안(Hinton et al., 2006; Bengio et al., 2007)되었고, 지금 까지 이미지 변환, 복구, 자료의 압축코드화, 음성 특징 스펙트럼 추출 등 많은 분야에 활용되고 있다.

### 2.3. 다점지구통계학(Multiple-point Geostatistics)

초기 저류층 모델은 코어자료, 로깅자료 등의 정적자료를 활용한 지구통계기법을 통해 생성할 수 있다. 이러한 지구 통계기법에는 한 번에 한 지점의 값을 예측하는 픽셀기반(pixel-based) 기술, 하나의 대상이나 패턴단위로 객체를 배정하는 객체기반(object-based) 기술과 두 방법의 단점을 극복한 다점지구통계기법이 있다. 픽셀기반 기법은 주어진 자료는 잘 보존하지만 패턴이 있는 지질구조를 모사하지 못하는 한계점이 있으며, 객체기반 기법은 패턴은 잘 나타내지만 주어진 정보를 잃어버리는 한계점이 있다(최종근, 2013). 반면 다점 기반 모델링은 트레이닝 이미지(TI; Training image)를 활용하여 주어진 정보를 보존하면서도 공간적 분포패턴을 모사할 수 있는 장점이 있다. TI는 저류층 물성값의 공간적 분포정보를 담고 있기 때문에 픽셀기반 기법의 베리오그램의 역할을 대체할 수 있으며, 이미지로부터 조건부확률을 직접 계산하여 저류층을 모델링한다. Figure 2.5는 TI로부터 조건부확률을 계산하는 예시이다. 저류층 모델의 가운데 지점을 예측하기 위해 주변의 암상을 만족하는 검색패턴을 구하였을 때 4개의 패턴을 구할 수 있고, 이 패턴으로부터 예측지점의 암상확률은 3/4의 확률로 사암, 1/4확률로 셰일성 사암이 되는 것을 계산할 수 있다. 이와 같이 다점지구통계기법은 채널 연결성을 잘 모사하면서 주어진 정적자료를 잘 보존할 수 있기 때문에 채널 저류층 모델링에 있어 많이 활용되고 있다.



**Figure 2.5 Training image and conditional probability in multiple-point geostatistics** (최종근, 2013).

### 2.3.1. SNESim (single normal equation simulation)

다점지구통계 기법이 제안된 초기에는 매 예측과정에서 전체 TI를 검색했기 때문에 느리고 비효율적이었다. 이를 보완하기 위해 TI에서 발생할 수 있는 이벤트를 세분화하고 빈도를 저장하여 한 번의 검색만으로 예측할 수 있는 탐색트리 개념이 제안되었으며, SNESim은 탐색트리 이용한 가장 기본적인 다점 지구통계기법이다. SNESim은 오직 범주 변수만을 다룰 수 있기 때문에 암상의 예측에는 유용하나 유체투과도와 같은 연속적인 변수값을 예측하는 데에는 어려움이 있다. SNESim에서는 채널 저류층의 지질 패턴을 효과적으로 모사하기 위해 템플릿에 사용된 지점수를 변경하거나 크기를 변경할 수 있다. 일반적으로 큰 템플릿은 큰 구조를 대략적으로 파악할 때 활용되며, 작은 템플릿은 모델링을 마무리할 때 활용된다. 따라서 여러 크기의 템플릿을 활용할 경우 정교한 모델링을 수행할 수 있으며, 이를 멀티그리드 시뮬레이션이라 한다(최종근, 2013).



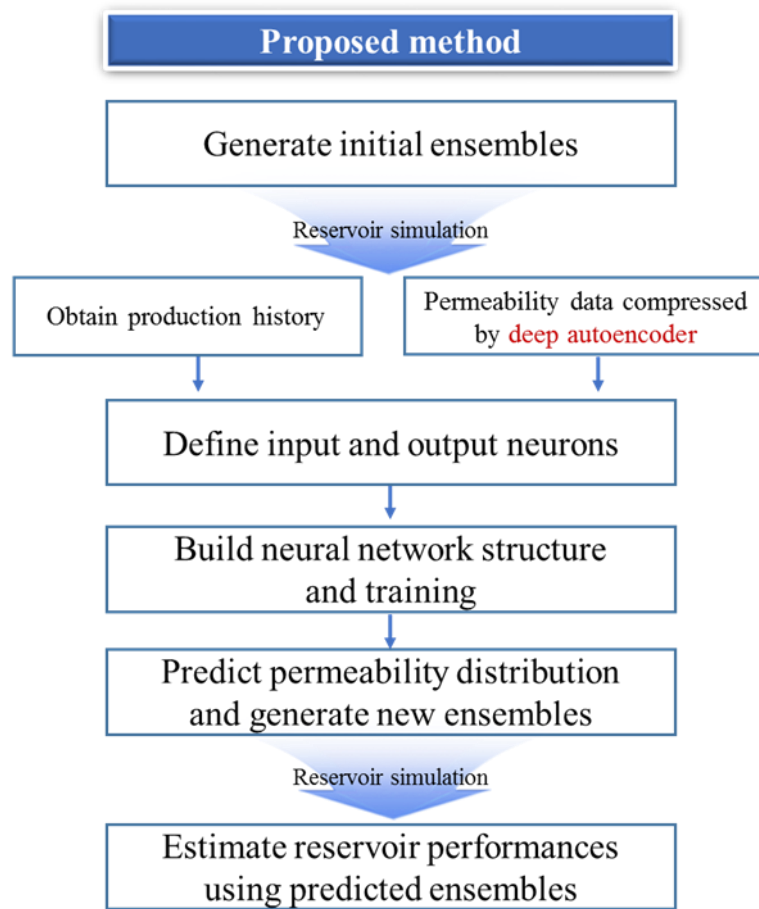
### 3. 연구 방법

이 논문에서는 저류층의 유체투과도 분포와 생산정의 미래생산추이를 예측하기 위하여 저류층 특성화를 수행하였다. 특성화를 수행하기 위해 TI로부터 생성된 다수의 저류층 모델인 앙상블을 신경망을 통해 학습시켰으며, 신경망으로부터 예측된 앙상블의 해집합을 활용하여 저류층 생산성의 불확실성을 평가하였다. 신경망을 통해 동적자료인 생산이력과 정적자료인 유체투과도 분포의 비선형관계를 학습하였다. 그리고 심층 오토인코더를 적용함으로써 유체투과도의 올바른 분포정보를 보존할 뿐만 아니라 연산속도를 증가 시킬 수 있었다. 이 장에서는 이를 위해 개발한 저류층 특성화 기법의 전체 순서도와 효과를 소개한다.

#### 3.1. 신경망을 이용한 앙상블 학습

Figure 3.1은 저류층 특성화를 위해 신경망을 이용한 앙상블 학습과 앙상블 모델의 유체투과도 값을 예측하는 과정을 나타낸 순서도이다. 먼저 복잡한 채널 경향을 보이는 트레이닝 이미지로부터 초기 앙상블을 생성한다. 학습을 위한 신경망구조의 구축에 앞서 입출력 자료를 설정해 준다. 입력자료의 설정은 3.1.1항에서 설명하였으며, 출력자료의 압축변환과정은 3.2항에서 설명한다. 신경망 학습을 통해 격자 별 유체투과도 값을 예측하고 학습을 반복하여 다수의 앙상블 해집합을 생성한 뒤 저류층의 성능과

불확실성을 평가하였다. 개발한 기법은 생산이력과 유체투과도의 비선형 관계를 예측하는 인공신경망 기법에 심층신경망의 오토인코더를 적용하여 불균질한 유체투과도 분포정보를 보존하면서 빠르게 연산할 수 있는 장점이 있다.



**Figure 3.1 Flow chart of the proposed method.**

### 3.1.1. 인공신경망 입출력 자료

참조필드로부터 얻은 생산이력(오일생산량과 물생산량)과 트레이닝 이미지로부터 생성된 앙상블의 생산이력과의 차이를 입력자료로 설정하였다. 각 앙상블의 생산이력이 입력자료로 설정한 것이 아니라 생산이력의 차이를 입력자료로 설정한 이유는 데이터 전처리 과정에서 발생할 수 있는 문제를 보완할 수 있기 때문이다. 신경망 학습에서 모든 입출력자료는 학습과정의 수치적 안정성을 위하여 자료 전처리 과정이 선행되어야 한다. 자료 전처리를 위한 기법으로 표준화(Standardizing)와 정규화(Normalizing)가 있으며 이러한 기법은 복잡한 입력자료들을 균일한 스케일로 재조정할 수 있다.

$$\textbf{Standardizing} : X_{new} = 2 * \frac{X - X_{min}}{X_{max} - X_{min}} - 1 \quad (3.1)$$

$$\textbf{Normalizing} : X_{new} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (3.2)$$

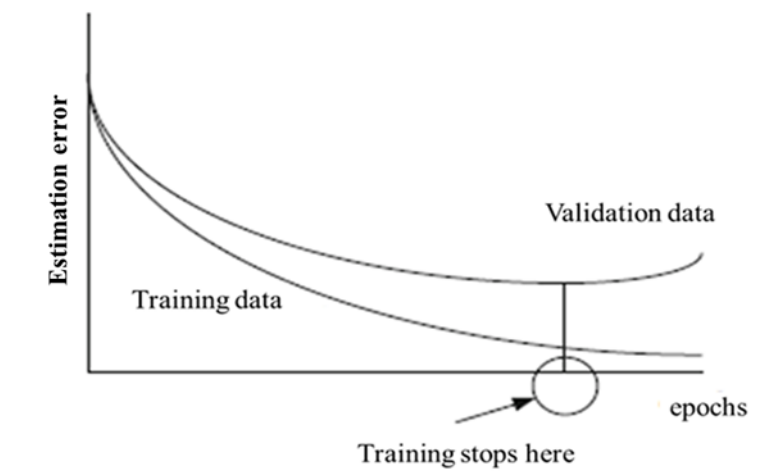
여기서  $X_{min}, X_{max}$  는 입출력자료의 최소값과 최대값을 나타내며,  $X$  는 입출력자료,  $X_{new}$  는 균일한 스케일로 조정된 입출력 값이다. 참조 필드의 오일 생산량이 1, 2, ..., 10, 11 이고 저류층 모델의 오일 생산량이 10, 20, ..., 100, 110 이라고 가정한다면, 생산이력의 차이가 아닌 생산이력자료 그대로가 입력자료로 설정될 경우 정규화 전처리 과정에서 동일한 스케일로 조정되어 신경망이 같은 입력자료로 판단하는 문제를 발생시킨다. 이러한 문제를 보완하기 위해 생산이력의 차이를 입력자료로 설정하였으며, 각각의 저류층 모델에 대해 신경망이 각각 다른 입력자료로 판단할 수 있게 된다. 입력자료를 생산이력의 차이로 나타내는 것은 거리기반기법과도 연관성이

있다. 거리기반기법은 비유사성을 나타내는 거리계산으로 저류층 거동과의 연관성을 예측하는데 효과적인 방법이다. 생산이력의 차이는 동적기반 거리라고 볼 수 있다.

동적자료인 생산이력을 이용하여 신경망 학습을 통해 정적 물성을 예측하는 연구에서 생산이력을 입력자료로 설정할 때 생산 시간과 생산량을 뉴런 벡터로서 설정하여 정적물성을 예측하였다(Ramgulam, 2006; Shahkarami et al., 2014). 이에 따라 동적자료로 활용되는 생산정의 수가 하나이며, 오일 생산량과 물생산량 자료가 있을 경우 시간과 함께 3개의 뉴런으로 입력 층이 구성된다. 하지만 유체투과도 분포에 따른 생산이력의 비선형 관계를 학습하는데 있어 단 시간에 따른 생산이력으로 격자 전체의 유체투과도를 예측하는 것은 큰 불확실성이 발생할 수 있다. 유체투과도 분포가 변하지 않는다고 가정하였을 경우 격자 별 유체투과도 값은 각 생산정의 모든 시간의 생산이력에 관여하기 때문에 단 시간이 아닌 모든 시간의 생산이력이 입력층의 뉴런으로 설정되어야 한다. 그러므로 만약 2개의 생산정에서 100일 동안 10일 간격의 생산이력(오일생산량, 물생산량)을 입력자료로 설정한 경우 입력 층의 뉴런 수는 40개가 된다. 이와 같은 방식으로 입력자료를 설정하였으며, 출력자료는 격자 별 유체투과도 값을 자료 전처리 과정을 거쳐 설정해 줌으로써 주어진 생산이력과 격자 별 유체투과도와의 비선형 관계 예측을 위한 입출력 자료를 설정하였다.

### 3.1.2. 인공신경망 구축 및 학습

신경망 학습에 앞서 입출력 자료 쌍은 일반적으로 세 가지로 나누어진다. 첫 번째 자료 쌍은 학습자료 쌍(training set)으로 자료를 이용하여 가중치를 변화시켜가며 학습을 하기 위한 자료 쌍이다. 두 번째 자료 쌍은 검증자료 쌍(validation set)으로 학습이 타당하게 진행되고 있는지를 판별하기 위한 자료 쌍이다. 학습이 진행되는 동안 검증자료 쌍의 오차가 지속적으로 관측되며, 초기에는 학습자료 쌍과 마찬가지로 오차가 감소하다가 신경망에서 과적합(overfit)이 발생하였을 때 오차가 급격히 증가한다. Figure 3.2는 신경망 네트워크의 종료시점을 결정하는 과정을 나타낸 것으로 네트워크의 검증자료쌍 오차가 최소가 되는 지점에서 학습을 종료 시키며, 이 시점에서의 가중치와 편향 값을 저장하여 사용하게 된다(Araghinejad, 2014).

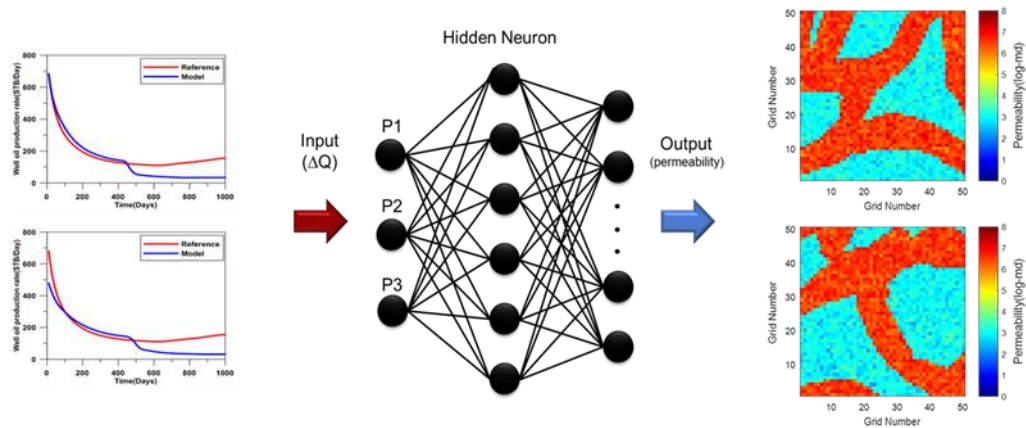


**Figure 3.2 Timing of training stop in data training and validation (Araghinejad, 2014).**

세 번째 자료 쌍은 테스트자료 쌍(test set)으로 학습자료 쌍과 검증자료 쌍으로

사용되지 않으며, 신경망의 성능을 평가비교하기 위해 사용된다. 각각의 자료 쌍들은 학습 가능한 자료 쌍의 수에 따라 분배할 수 있으며, 본 연구에서는 전체 자료 쌍으로부터 학습자료 쌍 80%, 검증자료 쌍 10%, 테스트자료 쌍 10%로 분배하였다.

신경망으로는 역전파 알고리즘을 이용한 피드포워드 신경망을 사용하였다. 피드포워드 신경망은 인공지능 분야에서도 널리 적용 (Doraisamy et al., 1998)될 뿐만 아니라 다른 신경망보다 빠른 연산이 가능하며, 많은 자료 쌍을 가진 문제를 처리할 수 있는 장점이 있다(Centilmen et al., 1999). 최적의 신경망을 구축하기 위해서는 고려해야 할 사항으로 신경망 입출력자료의 구조, 은닉층의 수, 은닉층을 구성하는 뉴런의 수, 활성화 함수, 학습 알고리즘, 성능평가 함수 등이 있다. 각각의 요소들은 학습의 시행착오를 통해 최적 값을 구하여 신경망 학습에 이용하였으며, Figure 3.3은 생산이력과 유체투과도 분포의 비선형관계를 예측하기 위해 구축한 인공신경망의 구조이다.



**Figure 3.3 Structure of artificial neural network.**

### 3.1.3. 신경망 학습을 통한 유체투과도 분포 예측

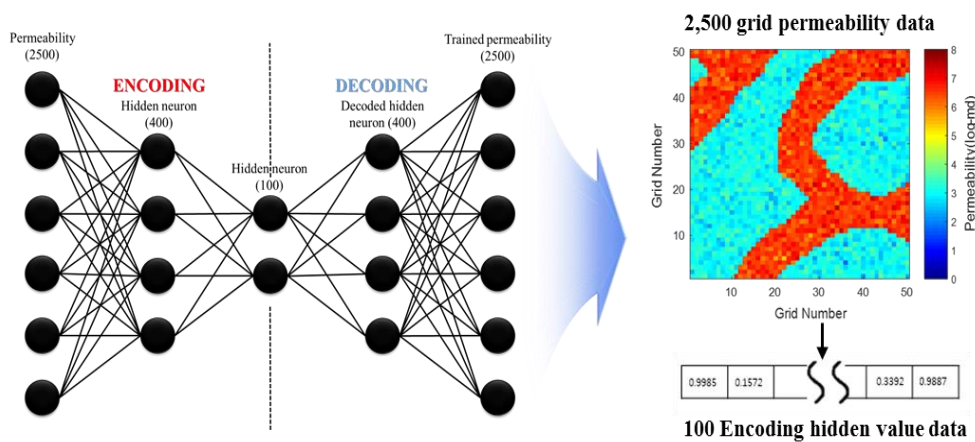
구축한 인공신경망을 이용하여 학습을 진행하였으며, 학습을 통해 참조모델의 유체투과도 분포를 예측하였다. 신경망의 입력자료는 참조모델과 학습할 앙상블들의 생산이력 차이값을 자료 전처리 과정을 통하여 표준화(Standardizing)하였다. 그러므로 만약 앙상블의 생산이력이 참조모델의 생산이력과 비슷할 경우 차이값은 0에 가깝게 되며, 표준화하게 되면 각 입력층의 뉴런 값이 -1에 가깝게 설정된다. 이를 이용하여 신경망 학습이 완료된 후 가상의 입력층 뉴런 값으로 -1값을 배정하여 입력층 뉴런값으로 설정하면 학습을 토대로 한 참조모델의 유체투과도 분포 예측할 수 있다. 예측되는 유체투과도의 분포는 학습자료, 검증자료, 테스트자료 쌍의 분배와 신경망 학습과정에서 계산되는 가중치와 편향에 따라 달라지기 때문에 100회의 독립적 학습과정을 통해 100개의 다른 유체투과도 분포를 가진 앙상블을 생성하였다. 이후 신경망으로부터 예측된 앙상블의 해집합을 활용하여 저류층 생산성의 불확실성을 평가하였다.

## 3.2. 심층 오토인코더의 적용

이 연구에서는 인공신경망을 이용하여 비선형관계를 가지는 동적자료인 생산이력과 정적물성인 유체투과도를 예측하는 기법을 제안하였다. 기존 연구의 균질한 유체투과도를 가진다는 가정과는 다르게 모든 격자의 비균질한 유체투과도를 예측하였다. 하지만 인공신경망 학습과정에서 출력자료가 50x50

격자모델의 유체투과도 값이기 때문에 2,500개의 출력층 뉴런이 필요하다. 일반적인 인공신경망 문제의 경우 출력층의 뉴런 수는 수 개에서 수십 개에 불과하지만 이 연구에서는 수천 개의 뉴런을 예측하는 문제로 바뀌면서 연산속도가 급격하게 느려지고, 예측 성능도 감소하게 된다.

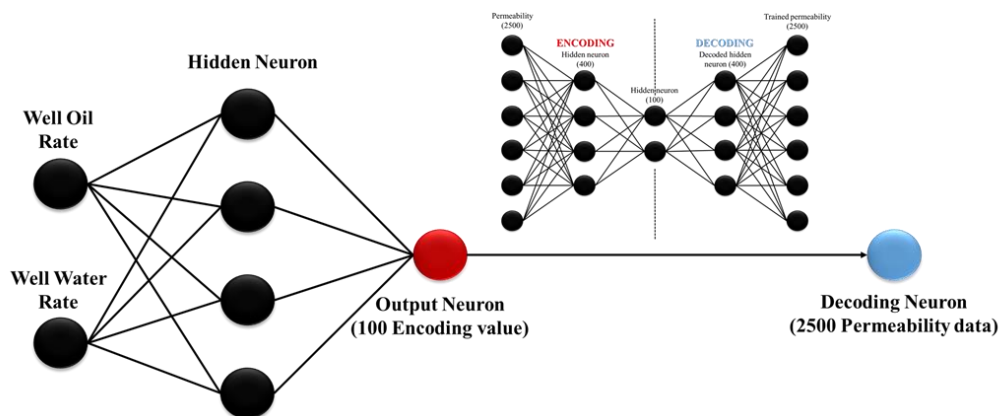
이와 같은 문제점을 보완하기 위해서 심층 오토인코더를 출력자료의 압축변환을 위해 적용하였다. 심층 오토인코더를 이용하면 2,500개의 격자 별 유체투과도 자료를 원하는 수의 자료로 압축변환 할 수 있으며, 이 연구에서는 100개의 자료로 압축하였다. Figure 3.4는 심층 오토인코더를 이용하여 2500개의 자료를 100개의 자료로 압축하는 과정이다. 출력층의 뉴런 수가 1/25로 감소하면서 신경망의 연산속도가 빨라졌으며, 압축변환 과정에서 심층 오토인코더가 유체투과도의 분포 경향을 학습하였기 때문에 신경망을 통해 예측하는 유체투과도의 분포 경향을 참조모델과 유사하게 모사할 수 있었다.



**Figure 3.4 Application of deep autoencoder for transforming 2,500 grid permeability data to 100 hidden value data.**



앞서 구축한 인공신경망 구조에 심층 오토인코더를 결합하여 출력자료를 변환 적용하였으며, 이 연구에서 개발한 신경망 기법의 구조는 Figure 3.5와 같다. 저류층 모델과 참조모델의 각 생산정의 오일생산량과 물 생산량의 생산이력 차이를 입력자료를 설정하였으며, 심층 오토인코더를 이용해 압축 변환한 100개의 새로운 뉴런 값을 출력자료로 설정하였다. 이 같은 입출력자료를 가지고 신경망 학습을 수행하였으며, 참조모델의 유체투과도 분포를 예측하였다. 신경망 학습을 토대로 예측한 결과는 기존 인공신경망 기법에서는 2,500개의 유체투과도를 직접 예측하지만, 개발한 신경망 기법에서는 심층 오토인코더로 변환 된 100개의 자료값을 예측한다. 예측 된 100개의 자료값은 앞서 압축변환을 위해 사용한 심층 오토인코더를 이용하여 해독과정(Decoding)을 수행하고 그 결과 2,500개의 유체투과도 값을 예측할 수 있다.



**Figure 3.5 Proposed neural network structure with deep autoencoder.**

## 4. 연구 결과

개발한 신경망 기법의 성능을 검증하기 위하여 채널저류층의 특성화 문제에 적용해 보았다. 생산정의 오일생산량과 물생산량을 동적자료로 이용하였으며, 격자별 유체투과도 분포 예측을 통해 특성화를 수행하였다. 저류층 특성화를 수행하기에 앞서 트레이닝이미지로부터 초기 앙상블을 생성하기 위해 스탠포드 대학에서 개발한 SGeMS(Stanford geostatistical modeling software)를 사용하였다. 그리고 각 생산정의 생산이력을 얻기 위하여 Schlumberger사의 ECLIPSE100을 저류층 시뮬레이터로 사용하였다. 신경망 기반기법은 MATLAB을 기반으로 수행되었다. 채널저류층 특성화 연구를 진행하며 사용한 가정은 아래와 같다.

- 채널저류층은 사암과 셰일성 사암 두 암종으로 이루어져 있으며 두 암상은 각각 정규분포의 형태를 보이는 이봉분포를 따른다.
- 사암의 유체투과도 200~1500 mD, 셰일성 사암의 유체투과도는 5~50 mD이다.
- 오일과 물만 존재하는 2상유동이다.
- 코어자료로 5개 지점(4개-사암, 1개-셰일)의 암상자료가 있다.
- 3개의 생산정과 1개의 주입정이 정해진 운영조건에 의해 운영된다.
- 10일 간격으로 관측하여 300일까지 생산이력을 이용하여 학습하고 이후 1000일까지의 생산추이를 예측한다.
- 생산정은 워터컷에 상관없이 주어진 기간 동안 생산을 지속한다.

4.1 절에서는 인공신경망 기법과 심층 오토인코더를 적용한 신경망 기법의 예측성능을 검증하기 위하여 참조 모델이 포함된 학습에 대한 성능을 평가하였다. 4.2 절에서는 참조모델이 포함되지 않은 학습을 통해 각 신경망 기법의 채널저류층 특성화를 위한 적용 가능성을 평가하였으며, 4.3 절에서는 기존 인공신경망 기법을 이용한 저류층 특성화 연구에서 사용한 균질한 유체투과도를 가지는 저류층 모델과의 비교를 통해 개발한 신경망 기법의 우수성을 검증하였다.

## 4.1. 참조모델 구축과 신경망 기법 검증

### 4.1.1. 참조모델

채널저류층 참조모델과 초기 양상블은 다점지구통계기법으로 생성하였다. 우선 TI를 생성하기 위하여 SGeMS의 TIGenerator 모듈을 사용하였으며, 사인곡선의 형태를 가지는 사암연결성을 가정하여 생성하였다. Figure 4.1은 생성된 TI를 나타내면 Table 4.1은 TI를 생성하기 위한 채널의 폭, 길이, 방향성 등에 관한 인자를 설정한 것이다.

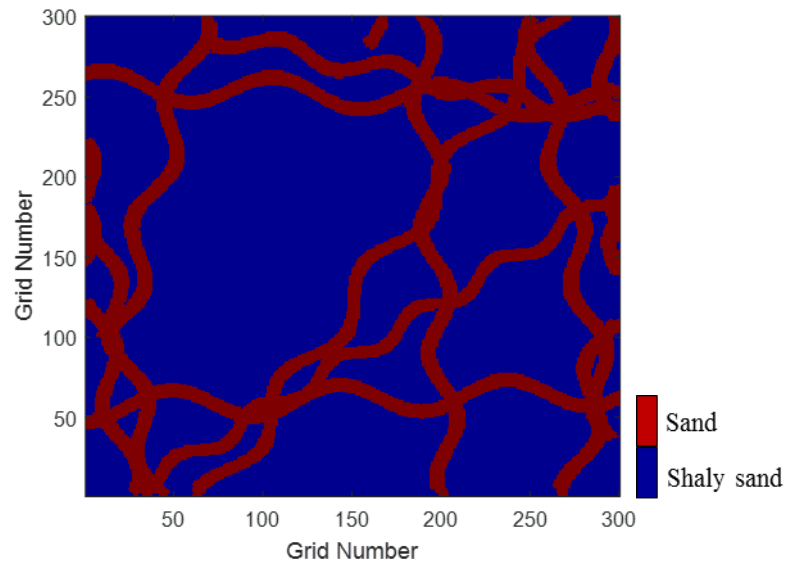
TI생성이 완료된 후, 이 TI와 SGeMS에서 제공하는 SNESim 모듈을 이용하면 초기양상블을 생성할 수 있다. SNESim모듈을 이용하기 위해서는 5개의 코어자료를 사용하였으며, Table 4.2는 SNESim 시뮬레이션 인자, Table 4.3은 4개 생산정의 암상자료와 운영조건을 나타낸다. Figure 4.2a는 시뮬레이션 결과 생성된 50x50 크기의 초기 양상블 중 선정한 참조필드로

0과 1로 대표되는 셰일성 사암(0)과 사암(1)으로 표현된다. Figure 4.2b는 SNESim으로 생성한 참조모델에 유체투과도 값을 배정하여 로그스케일로 분포를 나타낸 것이다.

참조모델은 중앙에 주입정을 두고 주입정을 기준으로 서북방향, 동북방향, 남쪽방향으로 각각 하나의 생산정이 둘러싸는 패턴을 가진다. 참조필드와 초기 앙상블에 적용할 저류층 물성은 Table 4.4와 같이 설정하였으며, 정해진 운영조건에 따라 생산하는 가정하에 저류층 시뮬레이션을 수행하였다.

**Table 4.1 Parameters for Training image**

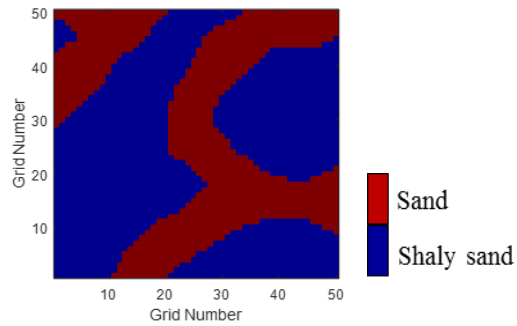
Parameters(unitless)	Values
Grid system(number of x, y, z)	300 x 300 x 1
Geobody type	Sinusoid
Length	Uniform distribution(mean=1000)
Width	Uniform distribution(mean=9)
Orientation	Triangular distribution (min=30, mode=180, max=270)
Amplitude	Triangular distribution (min=5, mode=8, max=10)
Wave Length	Triangular distribution (min=50, mode=80, max=100)
Sand ratio	0.25



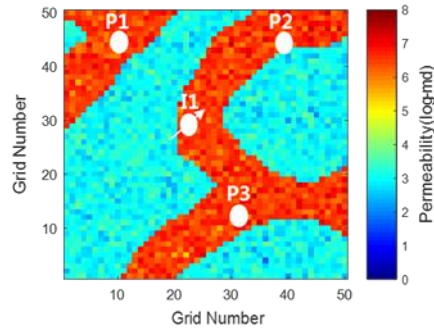
**Figure 4.1 Training Image.**

**Table 4.2 Parameters for SNESim simulation**

Parameters		Values
Grid system(number of x, y, z)		50 x 50 x 1
Target marginal distribution		Sand=0.6, Shaly sand=0.4
Search	Nodes	60
Template	Ranges	min=30, med=30, max=30



(a)



(b)

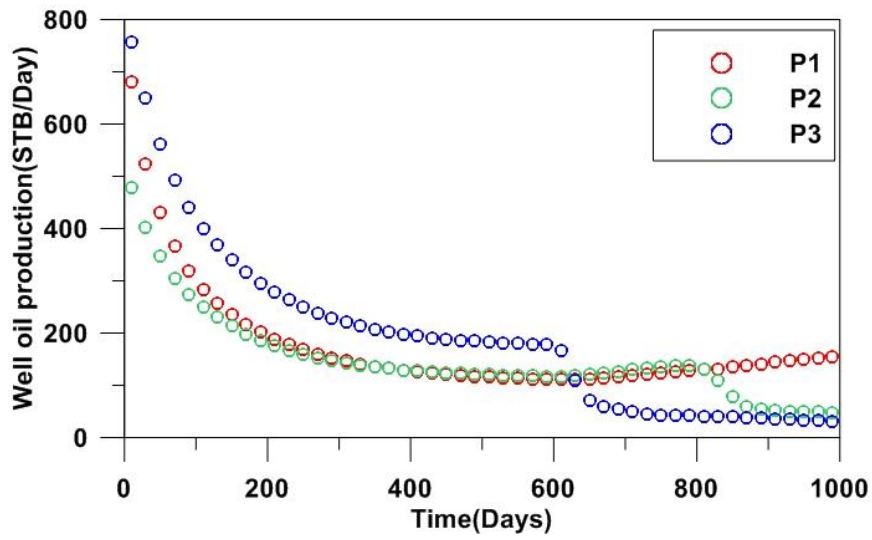
**Figure 4.2 Reference field: (a) realization of SNESim (b) log-permeability distribution.**

**Table 4.3 Well information and boundary conditions of the reference field and initial ensembles**

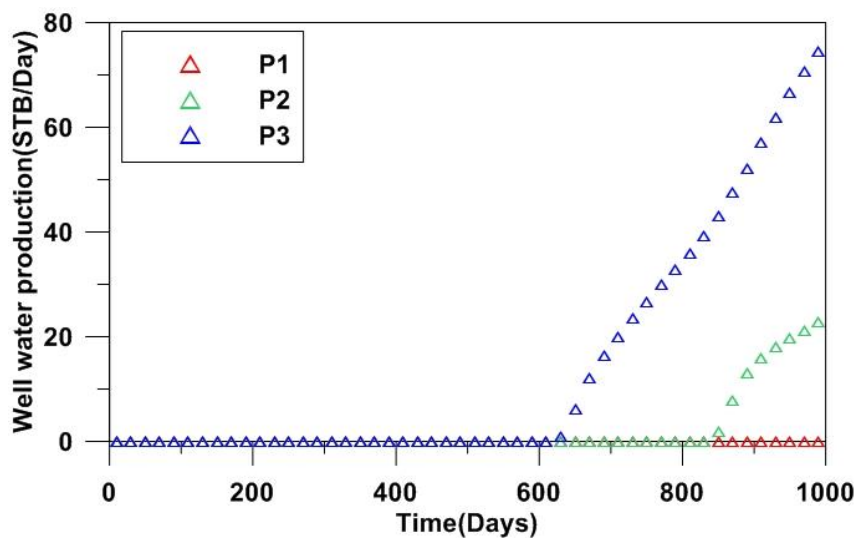
Well name	Location(X,Y)	Control mode, psia or STB/day	Facies
P1	(8, 43)	300	Sand
P2	(43, 43)	300	Sand
P3	(30, 10)	300	Sand
I1	(20, 30)	400	Sand

**Table 4.4 Petrophysical parameters for reference field and initial ensembles**

<b>Parameters</b>		<b>Values</b>
Grid system(number of x, y, z)		50 x 50 x 1
Width, ft		1,500
Length, ft		1,500
Thickness, ft		40
Porosity, fraction		0.3
Initial pressure, psia		2,000
Permeability, mD	Sand	200~1,500
	Shaly sand	5~50
Formation volume factor, rb/STB	Water	1 at 2,000 psig
	Oil	1.012 at 0 psig
		1.011 at 1,000 psig
		1.01 at 2,000 psig
Density of fluid, lb/ ft <sup>3</sup>	Water	48.6
	Oil	62.3
Viscosity of fluid, cp	Water	1
	Oil	3
Compressibility, psi <sup>-1</sup>	Water	5.0E-07 at 2,000 psig
	Rock	3.0E-05 at 2,000 psig



(a)



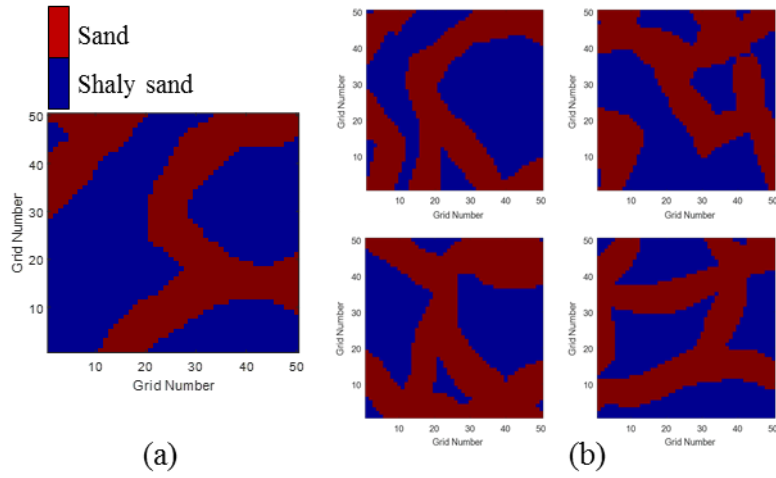
(b)

Figure 4.3 Well daily production of the reference field: (a) Oil rate (b) Water rate.



#### 4.1.2. 참조모델을 포함한 인공신경망의 학습성능

신경망 기법을 통해 채널저류층을 특성화하기 앞서 각 신경망 기법의 학습성능을 평가하였다. 이를 위하여 트레이닝 이미지로부터 생성된 초기 앙상블을 학습하였으며, Figure 4.4는 채널의 너비가 9인 트레이닝 이미지로부터 참조모델과 초기 앙상블을 생성한 것을 나타낸다. 학습성능의 평가를 위해 참조모델을 포함한 총 201개의 앙상블을 학습하였으며, Table 4.5는 인공신경망을 구축하기 위한 입출력변수를 나타낸다. 신경망의 입력자료는 참조모델과 초기 앙상블들의 3개 생산정의 생산이력 (오일생산량, 물생산량)차이 값이다. 10일 간격으로 300일간의 생산이력 차이를 입력층의 뉴런으로 설정하였으며, 이에 따라 입력층에는 오일 생산이력 차이 값 90개와 물 생산이력 차이 값 90개, 총 180개의 입력 뉴런이 설정되었다. 그리고 출력자료는 앙상블들의 격자 별 유체투과도 값이 설정 되었으며, 이에 따라 출력층에는 2,500개의 뉴런이 설정되었다. 입력층과 출력층의 비선형관계를 예측하기 위한 은닉층의 뉴런 수는 일반적으로 입출력층의 뉴런 수보다 많아야 하지만, 뉴런의 수가 증가함에 따라 연산성능의 증가보다 연산속도의 감소로 인한 효율성 감소가 크다. 이에 따라 은닉층의 뉴런 수는 3,000개로 설정하였다. 신경망의 학습알고리즘으로는 MATLAB의 신경망 알고리즘으로 변화율 기반의 역전파 알고리즘(Traindx, Gradient descent with momentum and adaptive learning rate backpropagation)을 사용하였다.



**Figure 4.4 Ensembles for neural network training: (a) reference (b) initial ensembles.**

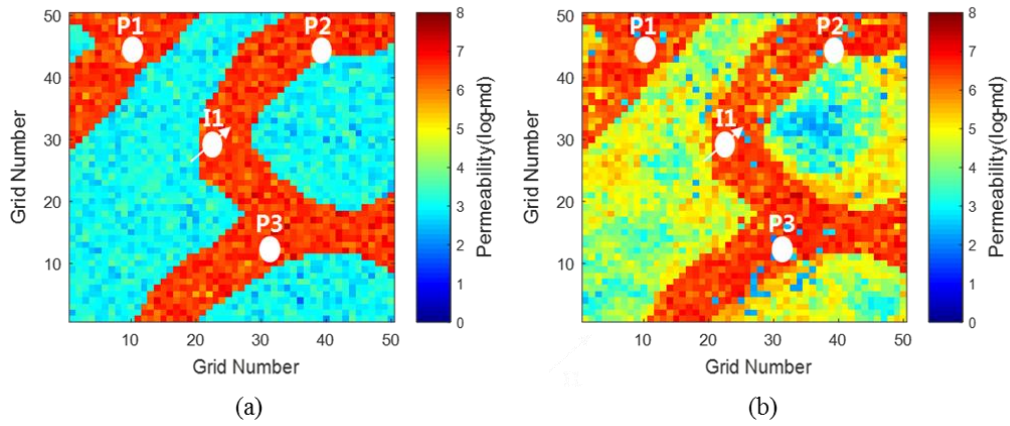
**Table 4.5 Neuron descriptions for the application of artificial neural network**

Description of the neuron		Number of neurons and algorithm
Input neuron	Well oil rate	90
	Well water rate	90
Hidden neuron		3,000
Output neuron	Grid permeability	2,500
Train algorithm		Traingdx <sup>†</sup>

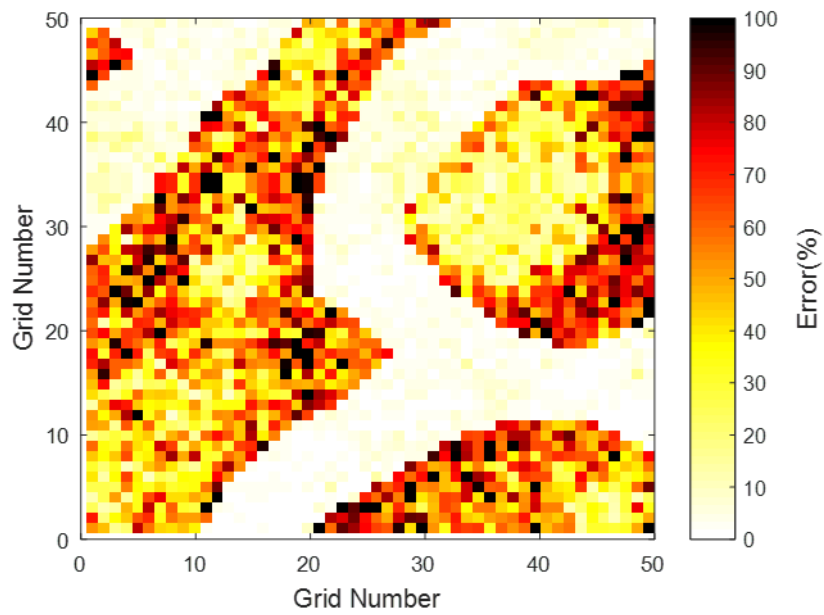
<sup>†</sup> Gradient descent with momentum and adaptive learning rate backpropagation

Figure 4.5는 참조모델이 포함된 초기 앙상블들을 인공신경망을 통해 학습하여 참조모델의 유체투과도 값을 예측한 결과이다. 참조모델이 포함된 학습이었기 때문에 채널연결성은 매우 근접하게 예측하였다. 하지만 인공신경망 기법을 적용할 경우 유체투과도 분포예측은 양호하지 못하였다. Figure 4.6은 예측한 100개의 앙상블들의 격자 별 유체투과도 오차를 평균하여 격자지도로 나타낸 것이며, 예측과정에서 사암지역의 유체투과도는 양호하게 예측하였지만, 셰일성 사암지역의 유체투과도 예측 오차율은 대부분의 격자에서 50%를 넘어서면서 예측에 실패한 것을 확인할 수 있다. 또한 Figure 4.7과 같이 히스토그램으로 나타내었을 때 낮은 로그유체투과도 값(2~4)은 대부분 예측하지 못하였다.

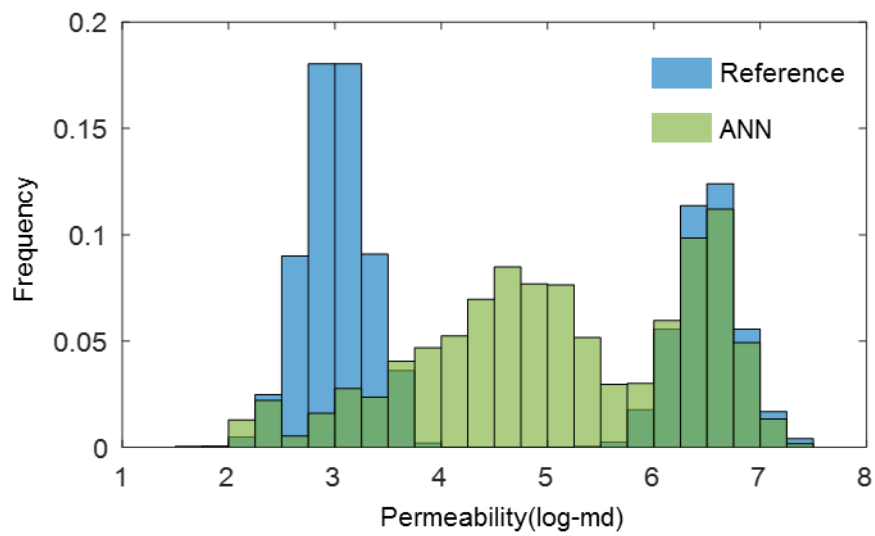
이로 인한 유체투과도 분포 예측의 실패는 미래생산량 추이 예측에도 영향을 주었다. Figure 4.8은 인공신경망 기법으로 예측한 100개의 앙상블들의 1,000일간 오일생산량과 물생산량 예측 결과이다. 300일간의 생산이력을 학습하여 추후 700일간의 생산량을 예측한 결과 물돌파가 발생하지 않은 1번 생산정에서의 생산추이는 오일과 물 생산 모두 양호하게 예측하였지만, 2번 생산정에서는 예측한 앙상블들은 물돌파를 예측하지 못하면서 참조모델의 생산추이와 다른 양상을 보였으며, 3번 생산정에서는 물돌파 시점을 참조모델보다 늦게 예측하여 생산추이 예측에 실패하였다. 이러한 결과, 인공신경망 기법의 성능평가에 있어서 학습을 통한 채널연결성예측은 양호하였으나, 유체투과도 분포에서 낮은 지역의 유체투과도 값 예측에 실패하면서 이로 인한 생산정들의 미래생산추이 예측성능이 낮은 것을 확인하였다.



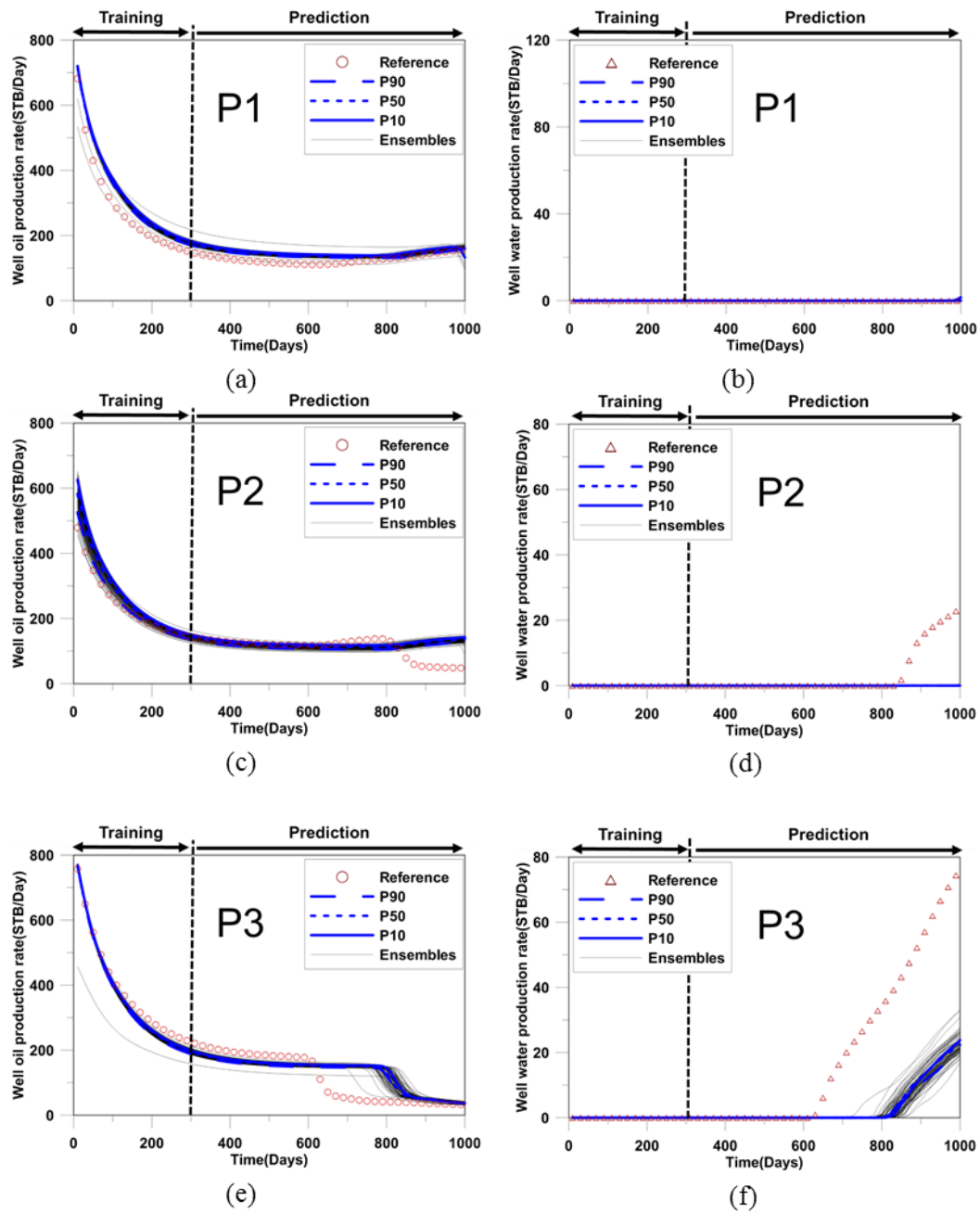
**Figure 4.5 Log permeability distribution: (a) reference (b) predicted ensemble(P50) by ANN in case of including the reference.**



**Figure 4.6 Averaged error of permeability predicted by ANN in case of including the reference.**



**Figure 4.7 Histogram of reference and predicted ensemble(P50) by ANN in case of including the reference.**



**Figure 4.8 Well production rate of predicted ensembles using ANN with reference: (a) oil rate of P1 (b) water rate of P1 (c) oil rate of P2 (d) water rate of P2 (e) oil rate of P3 (f) water rate of P3.**

#### 4.1.3. 참조모델을 포함한 경우 개발한 신경망의 학습성능

4.1.2절에서 참조모델을 포함한 학습에 인공신경망 기법을 적용하였을 경우 학습성능에 대해 평가하였다. 평가 결과 채널 연결성은 양호하게 예측하였지만, 유체투과도 분포 예측에 실패하였다. 본 절에서는 인공신경망 기법을 적용하면서 발생한 문제점들을 보완하기 위하여 심층 오토인코더를 신경망기법에 적용하였다.

학습에 사용된 양상블은 4.1.2절과 같았으며, Table 4.6은 오토인코더가 적용된 신경망을 구축하기 위한 입출력 변수를 나타낸다. 입력층의 뉴런은 기존 신경망과 같이 180개로 동일하며, 은닉층과 출력층의 뉴런수가 오토인코더의 압축변환에 따라 달라지게 되었다. 출력자료인 격자 별 유체투과도 값 2,500개의 자료는 오토인코더 학습에 따라 첫번째 인코더 층에서 400개, 두번째 인코더 층에서 100개의 은닉 값으로 압축되었다. 이와 같이 압축 변환된 100개의 자료값이 신경망의 출력층의 뉴런으로 설정되었으며, 은닉층도 필요한 뉴런의 수가 감소함에 따라 다층구조를 가질 수 있었으며 첫번째 은닉층은 600개, 두번째 은닉층은 300개의 뉴런수가 배정되었다. 오토인코더의 학습알고리즘은 MATLAB의 신경망 모듈 알고리즘으로 공액 변화율 기반의 역전파 알고리즘(Trainscg, Scaled conjugate gradient backpropagation)이다.

**Table 4.6 Neuron descriptions for the application of proposed neural network using autoencoder**

Description of the neuron		Number of neurons and algorithm
Encoding neuron	1 <sup>st</sup> Autoencoding	400
	2 <sup>nd</sup> Autoencoding	100
Hidden neuron[Multi-layer]		[600, 300]
Output neuron	Encoded Grid permeability	100
Train algorithm	Autoencoder	Trainscg <sup>†</sup>
	Neural network	Traingdx <sup>††</sup>

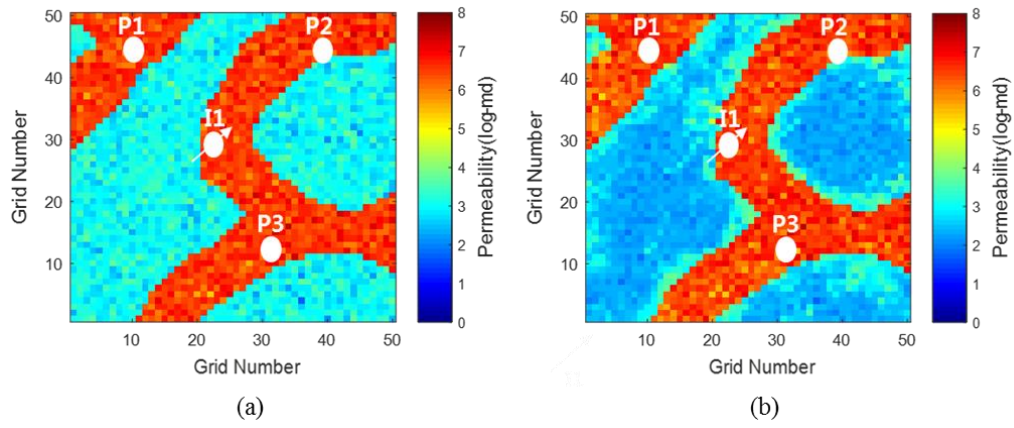
<sup>†</sup> Scaled conjugate gradient backpropagation

<sup>††</sup> Gradient descent with momentum and adaptive learning rate backpropagation

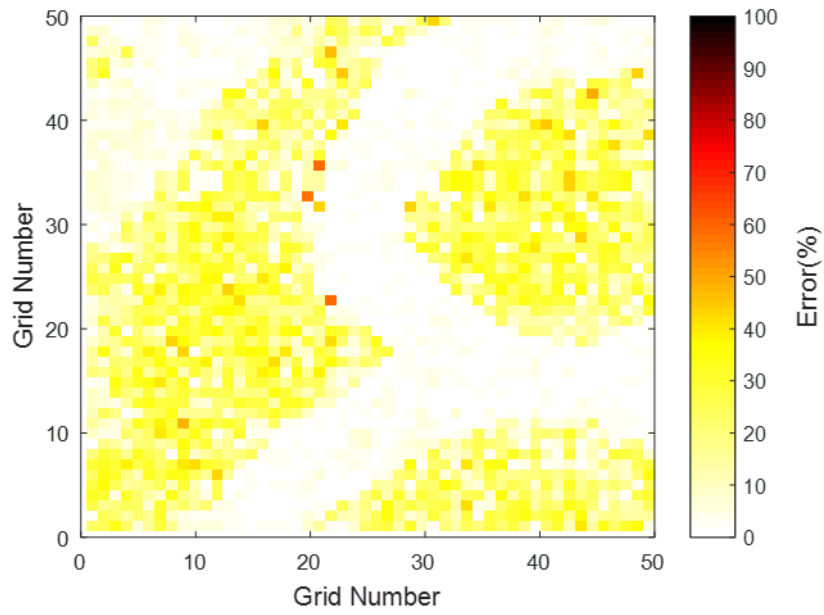


Figure 4.9는 참조모델이 포함된 초기 앙상블들을 심층 오토인코더를 적용한 신경망을 통해 학습하여 참조모델의 유체투과도 값을 예측한 결과이다. 인공신경망 기법을 적용한 경우와 같이 참조모델이 포함된 학습이었기 때문에 채널연결성은 매우 근접하게 예측하였다. 그리고 유체투과도 분포결과에서 확인할 수 있듯이 참조모델의 유체투과도 분포와 매우 근접하게 예측하였다. 인공신경망 기법을 적용할 경우 낮은 유체투과도 값 예측에 실패하여 Figure 4.6의 오차 격자지도에서와 같이 세일성 사암지역의 격자 별 오차율이 50%를 넘어선 반면, 개발한 신경망 기법은 Figure 4.10과 같이 오차 격자지도에서 세일성 사암지역의 격자별 오차율은 인공신경망 기법과 대비하여 낮은 것을 확인할 수 있었다. Figure 4.11은 개발한 신경망 기법으로 예측한 유체투과도를 히스토그램으로 나타낸 것이며, 높고 낮은 유체투과도 값 모두 매우 양호하게 예측한 것을 확인할 수 있다.

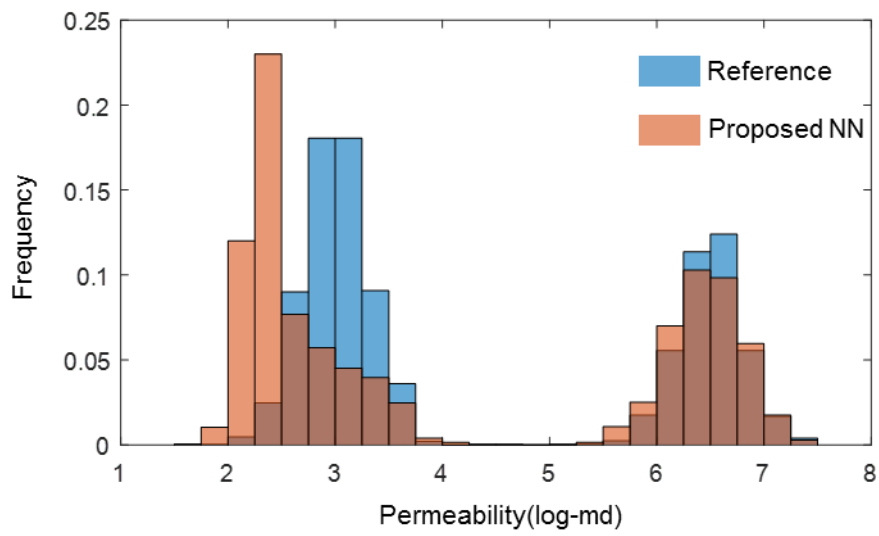
인공신경망 기법을 적용한 경우와 달리 심층 오토인코더를 이용하여 개발한 신경망 기법은 채널연결성뿐만 아니라 유체투과도 분포 예측을 참조모델과 근접하게 함으로써 미래생산량 추이 또한 참조모델과 유사하였다. Figure 4.12는 개발한 신경망 기법으로 예측한 100개의 앙상블들의 1,000일간 오일생산량과 물생산량 예측 결과이다. 1번, 2번, 3번 생산정 모두 물돌과 시점과 생산추이를 참조모델과 매우 근접하게 예측하였다. 이러한 결과를 통해 개발한 기법의 학습성능을 평가하였으며, 개발한 심층 오토인코더를 이용한 신경망 기법이 생산정의 생산이력과 채널연결성 및 유체투과도 분포의 비선형관계를 규명할 수 있다는 것을 검증하였다.



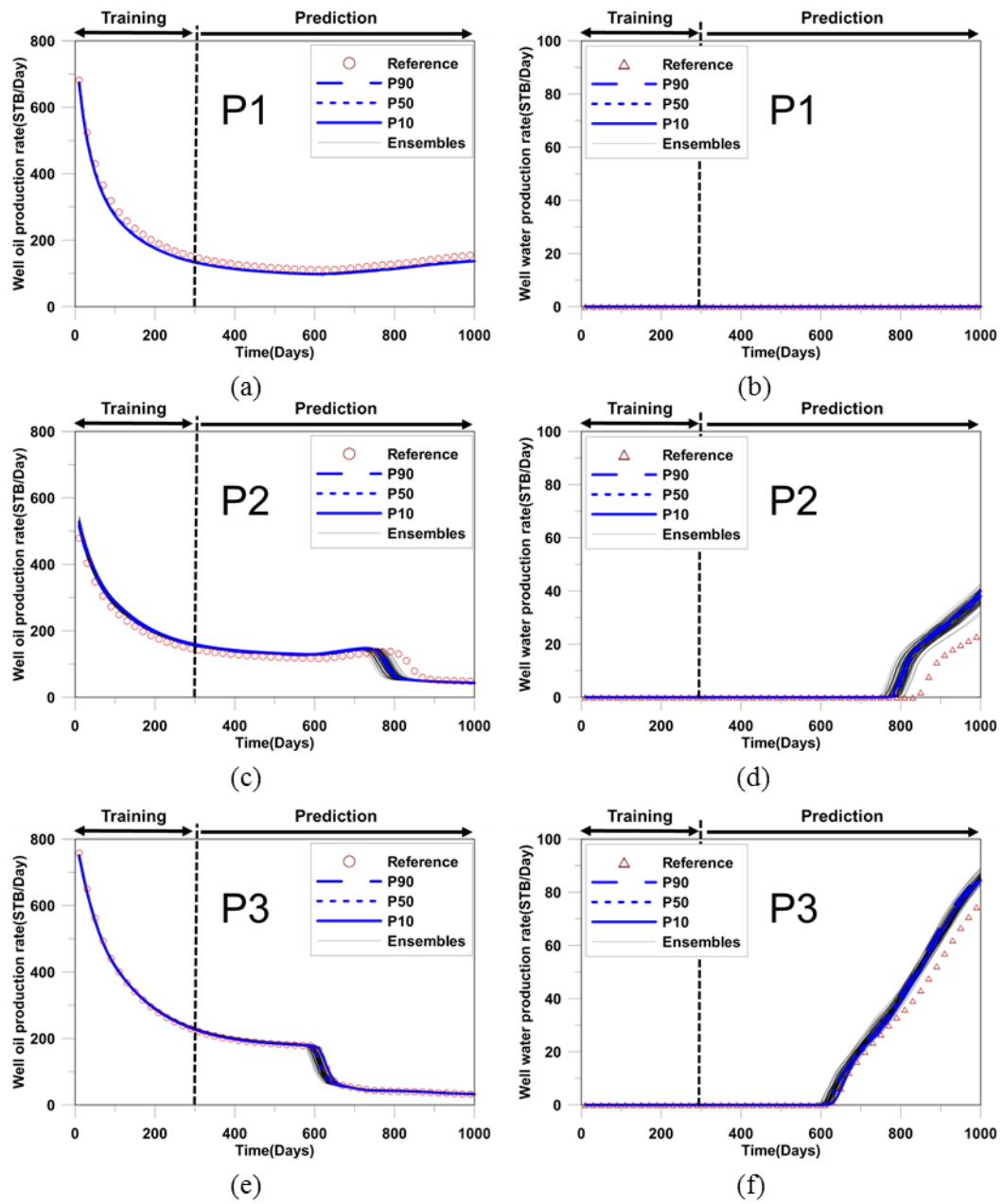
**Figure 4.9 Log permeability distribution: (a) reference (b) predicted ensemble(P50) by proposed NN in case of including the reference.**



**Figure 4.10 Averaged error of permeability predicted by proposed NN in case of including the reference.**



**Figure 4.11 Histogram of reference and predicted ensemble(P50) by proposed NN  
in case of including the reference.**



**Figure 4.12 Well production rate of predicted ensembles with reference using proposed NN:**  
**(a) oil rate of P1 (b) water rate of P1 (c) oil rate of P2 (d) water rate of P2 (e) oil rate of P3 (f)**  
**water rate of P3.**

## 4.2. 신경망 기법의 예측성능 평가

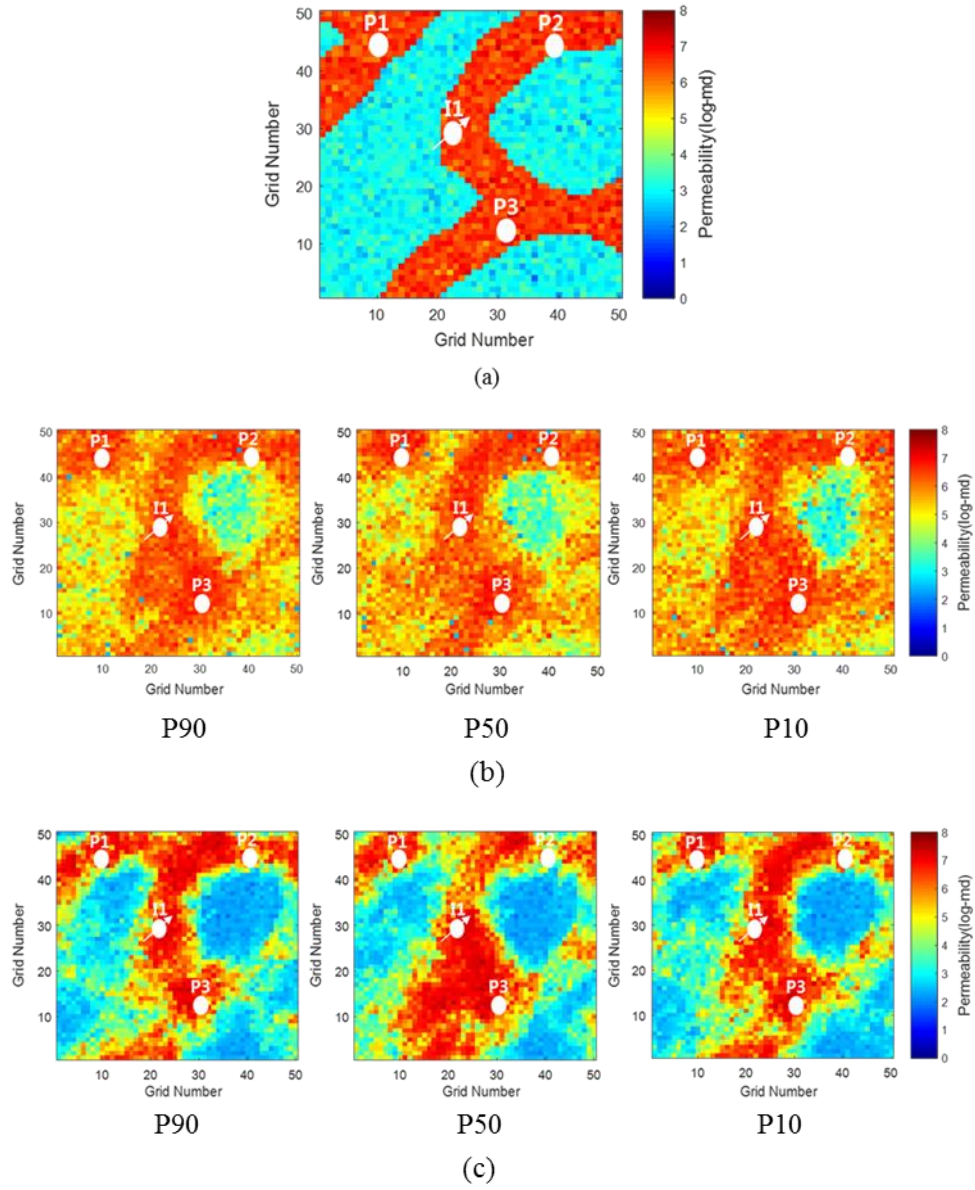
### 4.2.1. 인공신경망 기법과 개발한 신경망 기법의 채널저류층 특성화

4.1절에서는 참조모델을 설정하고, 기존의 인공신경망 기법과 개발한 신경망 기법의 성능을 검증하기 위해 참조모델을 포함한 학습을 수행하여 채널저류층을 특성화하였다. 이 절에서는 검증한 신경망 기법을 이용하여 채널연결성 및 유체투과도 분포의 예측을 통한 생산정의 미래생산추이를 예측 성능을 평가하기 위하여 참조모델이 포함되지 않은 경우에 대하여 학습을 수행하였다. 학습할 앙상블들에 참조모델이 포함되지 않았기 때문에 채널의 연결성 및 유체투과도의 분포는 초기 앙상블들을 바탕으로 생산정의 생산이력과 유체투과도의 분포의 비선형관계 학습성능에 따라 다르게 예측된다.

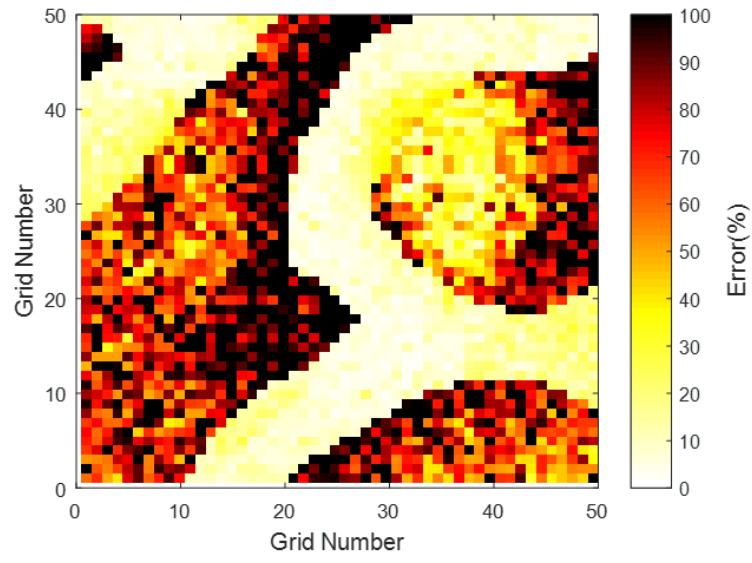
Figure 4.13b는 참조모델이 포함되지 않은 앙상블을 인공신경망 기법으로 학습하여 참조모델의 유체투과도 값을 예측한 결과이며, Figure 4.13c는 심층 오토인코더를 적용하여 개발한 신경망 기법으로 학습하여 예측한 유체투과도 값의 분포이다. 참조모델이 포함되지 않은 학습임에도 채널연결성은 유사하게 모사하였으며, 생산정 인근 배유구역(drainage area)의 유체투과도 값 또한 양호하게 예측하였다. 하지만 인공신경망 기법을 적용한 경우 Figure 4.14a의 오차 격자지도에서 나타난 것처럼 세일성 사암지역 대부분의 유체투과도 오차율이 50%를 넘어섰으며 사암과 세일성 사암의 경계지역에서는 오차율이 100%에 가까웠다. 또한 Figure 4.15a와 같이 낮은

로그유체투과도 범위(2~4)를 예측하지 못하고, 로그유체투과도 히스토그램이 5~6의 평균값을 가지는 정규분포 형태를 보이면서 유체투과도 값 예측에 실패한 것을 확인할 수 있다.

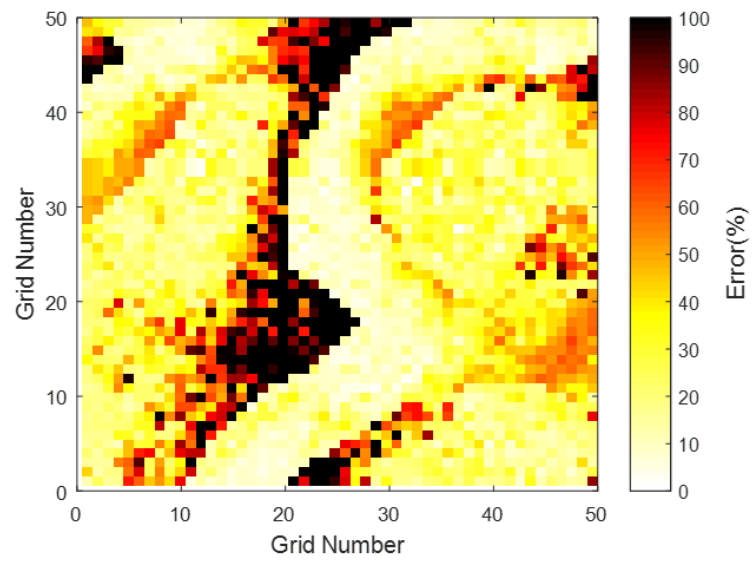
반면 심층 오토인코더를 적용한 신경망 기법은 Figure 4.14b의 오차 격자지도에서 나타난 것처럼 일부 사암과 셰일성 사암의 경계지역을 제외할 경우, 대부분의 유체투과도 오차율이 30%이하였으며 Figure 4.15b와 같이 낮은 유체투과도 값까지 포함 채널의 유체투과도 분포특성을 보존하였다. 개발한 신경망 기법은 참조모델이 포함되지 않아 참조모델의 유체투과도 분포를 사전에 학습하지 못했음에도 초기 앙상블들의 생산이력과 유체투과도의 비선형관계 학습을 통하여 유체투과도 분포를 예측하였다. 그 결과, 개발한 기법을 적용할 경우 학습된 앙상블들의 지질학적 불확실성으로 인해 일부 사암과 셰일성 사암의 경계지역에서 유체투과도 오차율이 50%를 넘어서면서 다소 예측성능이 떨어졌지만, 인공신경망 기법과 달리 전체적으로는 사암과 셰일성 사암의 유체투과도 분포특성을 보존한 것을 확인할 수 있다.



**Figure 4.13 Log permeability distribution of predicted ensembles: (a) reference (b) ANN (c) proposed NN in case of excluding the reference.**



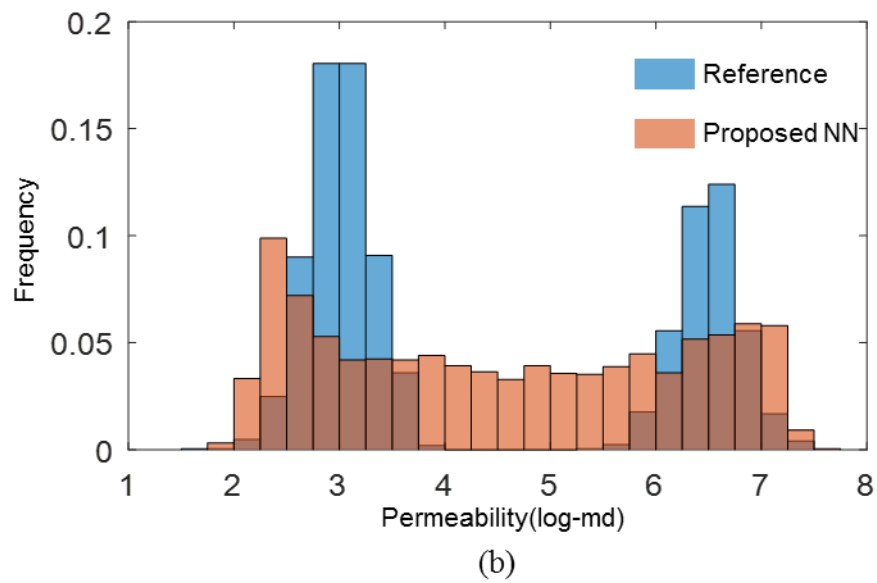
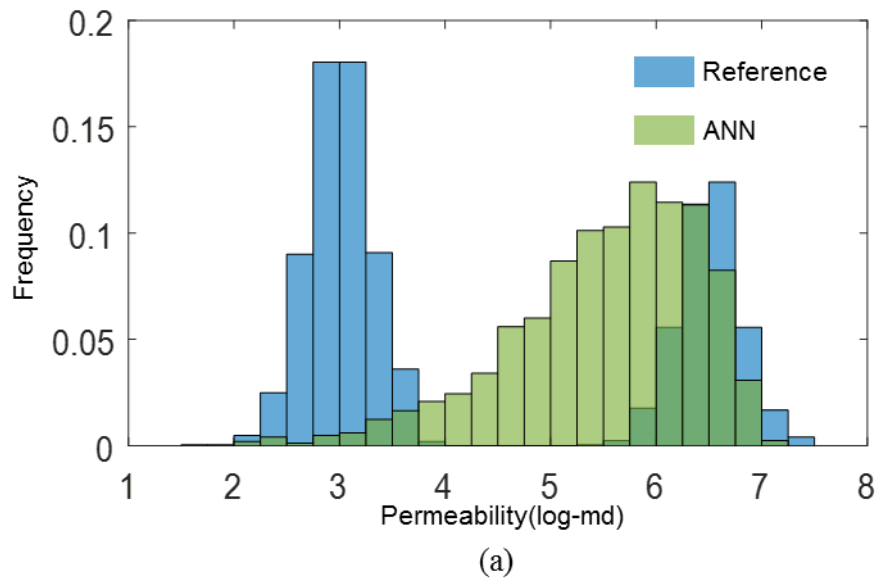
(a)



(b)

**Figure 4.14 In case of excluding the reference, averaged error of permeability predicted by: (a) ANN (b) proposed NN.**

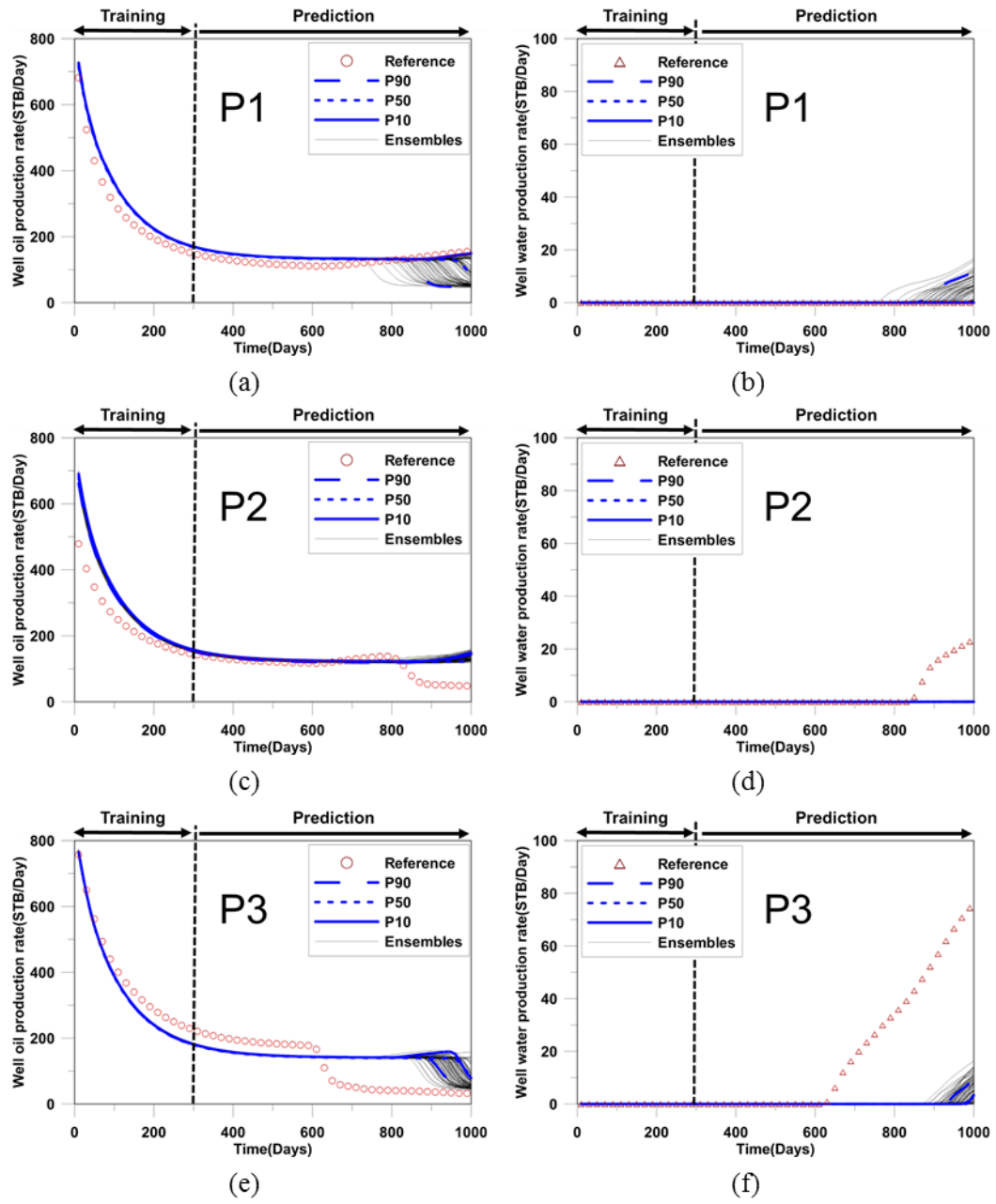




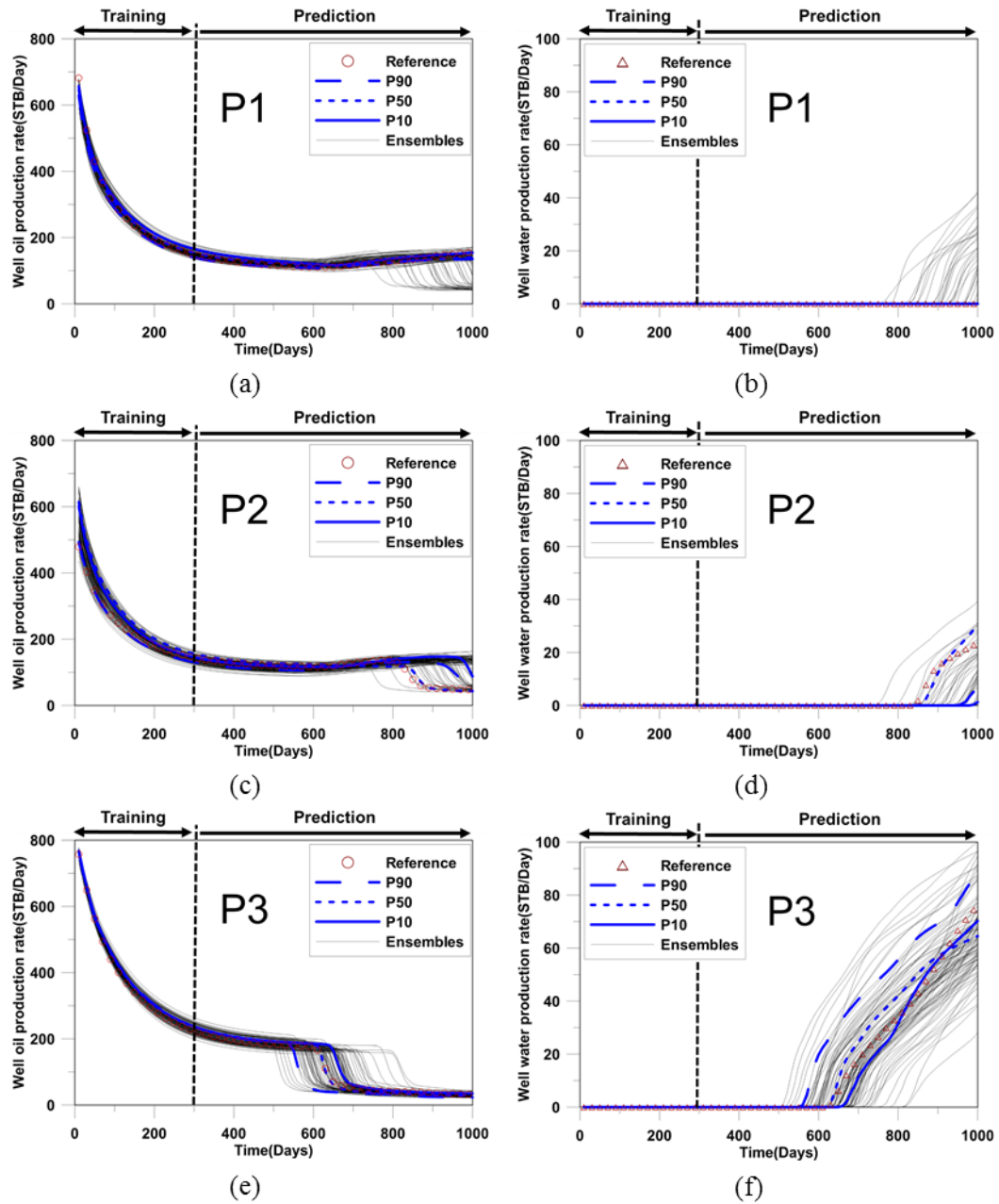
**Figure 4.15 In case of excluding the reference, histogram of reference and predicted ensemble(P50) by: (a) ANN (b) proposed NN.**

유체투과도 분포 예측은 생산정의 미래생산 추이 예측에 큰 영향을 주었다. Figure 4.16은 인공신경망 기법으로 학습하여 예측한 100개의 앙상블들의 1,000일간의 오일생산량과 물생산량 생산추이이다. 1번 생산정에서는 물돌파가 발생하지 않음에도 예측한 다수의 앙상블이 물돌파가 발생할 것이라 예측하였고, 2번과 3번 생산정에서 발생하는 물돌파의 시점을 제대로 예측하지 못하였다. 이는 참조모델과 대비하여 낮은 유체투과도 값 예측에 실패하고 저류층 격자의 유체투과도를 높게 예측하면서 주입정을 통한 물의 흐름이 사암으로 구성된 채널뿐만 아니라 셰일성 사암지역까지 발생하였기 때문이다.

반면 심층 오토인코더를 적용한 신경망 기법은 채널저류층의 유체투과도 분포특성을 보존하였기 때문에 미래생산량 예측에서도 우수한 성능을 보이는 것을 확인하였다. Figure 4.17은 개발한 신경망 기법으로 예측한 100개의 앙상블들의 미래생산량 예측 결과이다. 물돌파가 발생하지 않은 1번 생산정에 대해 대다수의 앙상블이 물돌파가 발생하지 않는다고 예측하였다. 하지만 2번 생산정에서는 P50을 비롯한 몇몇의 앙상블을 제외하고는 물돌파시점 예측에 실패하였다. 반면 3번 생산정에서는 참조모델의 생산이력이 불확실성에 포함 되었으며, P50의 경우 참조모델의 생산추이를 매우 근접하게 예측하는 등, 물돌파 발생시점을 신뢰할 수 있게 예측한 것을 확인할 수 있다.



**Figure 4.16 Well production rate of predicted ensembles without reference using ANN: (a) oil rate of P1 (b) water rate of P1 (c) oil rate of P2 (d) water rate of P2 (e) oil rate of P3 (f) water rate of P3.**



**Figure 4.17 Well production rate of predicted ensembles without reference using proposed NN: (a) oil rate of P1 (b) water rate of P1 (c) oil rate of P2 (d) water rate of P2 (e) oil rate of P3 (f) water rate of P3.**

#### 4.2.2. 인공신경망 기법과 개발한 신경망 기법의 비교

앞 절에서 인공신경망 기법을 적용한 경우와 심층 오토인코더를 이용한 신경망 기법을 적용한 경우에 대해 유체투과도 분포와 미래생산량 예측을 통한 채널저류층 특성화를 수행하였다. 그 결과, 인공신경망 기법을 적용한 경우 채널저류층 특성화에 실패하였지만, 오토인코더를 적용한 신경망 기법을 적용한 경우 생산이력과 유체투과도의 비선형관계를 규명함으로써 성공적으로 채널저류층을 특성화 할 수 있었다. Figure 4.18은 각 경우 생산정의 일일 오일 생산량에 대한 오차범위를 박스플롯(Box-plot)으로 나타낸 것이다. 오차 계산은 평균 절대오차율(MAPE; Mean Absolute Percentage Error)로 식(4.1)을 따른다. 참조모델이 포함 된 경우 앙상블들의 평균 절대오차율 중간값이 인공신경망은 32%, 개발한 신경망 기법은 8.6%였으며, 참조모델이 포함되지 않은 학습을 통한 예측성능으로 인공신경망은 45.9%, 개발한 기법은 15.6%의 오차율을 보였다.

$$MAPE = \frac{1}{n} \sum_{w=1}^n \left( \frac{100}{T} \sum_{t=1}^T \frac{|(WOPR_{ref,w}(t) - WOPR_{i,w}(t))|}{WOPR_{ref,w}(t)} \right) \quad (4.1)$$

식 (4.1)에서  $WOPR_{ref,w}$ : 참조필드의 일일 오일생산량

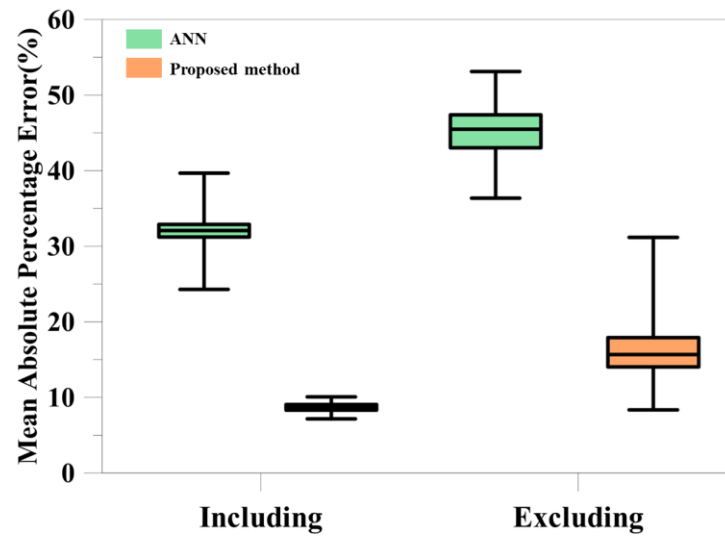
$WOPR_{i,w}$  : 저류층 모델의 일일 오일생산량

$n$  : 생산정 개수

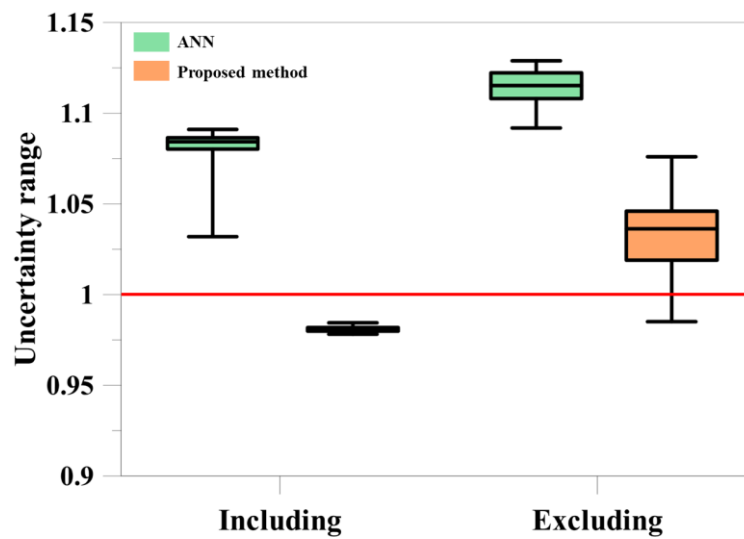
$T$  : 관측점 수

또한 1,000일간의 생산 동안 필드의 오일 누적생산량 불확실성을 비교해보았다. Figure 4.19는 각각 기법을 적용한 앙상블들의 오일 누적생산량을 박스플롯으로 나타낸 것이다. 참조모델의 누적생산량이 1이라고 보았을 때, 1값을 상자가 포함하거나 가까울수록 예측성능이 우수하다. 개발한 신경망 기법의 경우 누적생산량의 불확실성에 참조모델이 포함되어 있으며, 인공신경망 기법에 대비하여 참조모델의 값과 가까운 것을 확인할 수 있다.

심층 오토인코더를 적용한 신경망 기법은 미래생산추이 예측성능에서도 기존의 인공신경망과 대비하여 1/3수준의 오차율을 보였으며, 필드의 오일 누적생산량 불확실성 평가에서도 신뢰할 수 있는 결과를 도출하였다. 이와 같이 인공신경망 기법과 개발한 신경망 기법의 예측성능에서 차이가 발생하는 것은 채널저류층의 사암지역과 셰일성 사암지역의 유체투과도 값이 극단적으로 분포하는 특성으로 인하여 인공신경망이 비선형관계 학습을 위한 근사(fitting)과정에서 내삽(interpolation)하는 경향이 발생하였기 때문이다. 이로 인해 낮은 지역의 로그유체투과도(2.5~3.5)를 예측하지 못하고, 4~7사이의 연속적인 분포로 예측하였다. 반면 개발한 기법은 심층 오토인코더를 이용하여 2,500개의 자료를 100개의 자료값으로 변환하였고, 채널저류층의 극단적인 유체투과도 분포를 사전에 학습하였다. 이를 통해, 생산이력과 유체투과도의 비선형관계 학습을 위한 신경망학습에서 내삽 경향성이 발생하는 문제점을 보완하고 유체투과도의 분포정보를 보존할 수 있었다.



**Figure 4.18 Comparison to boxplot of daily well oil productions error between ANN and proposed NN in case of including and excluding the reference.**

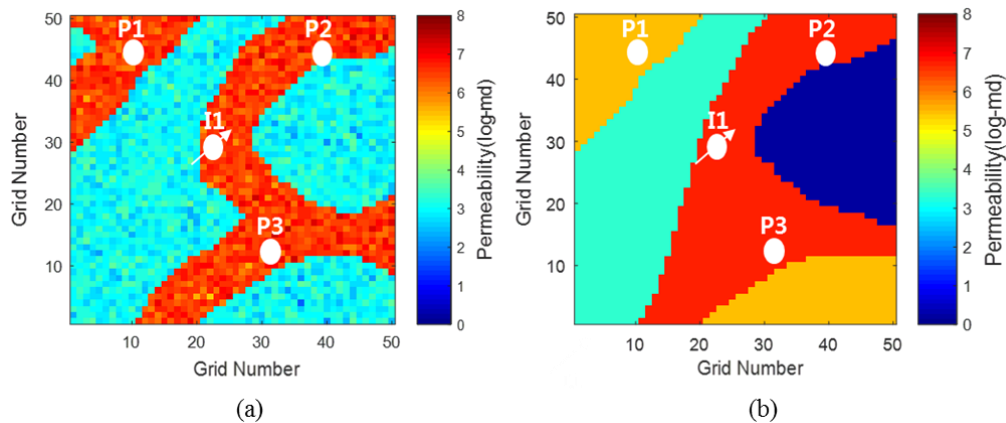


**Figure 4.19 Comparison to boxplot of uncertainty range of cumulative oil productions at 1,000 days between ANN and proposed NN in case of including and excluding the reference.**

### 4.3. 균질한 유체투과도 모델과의 비교

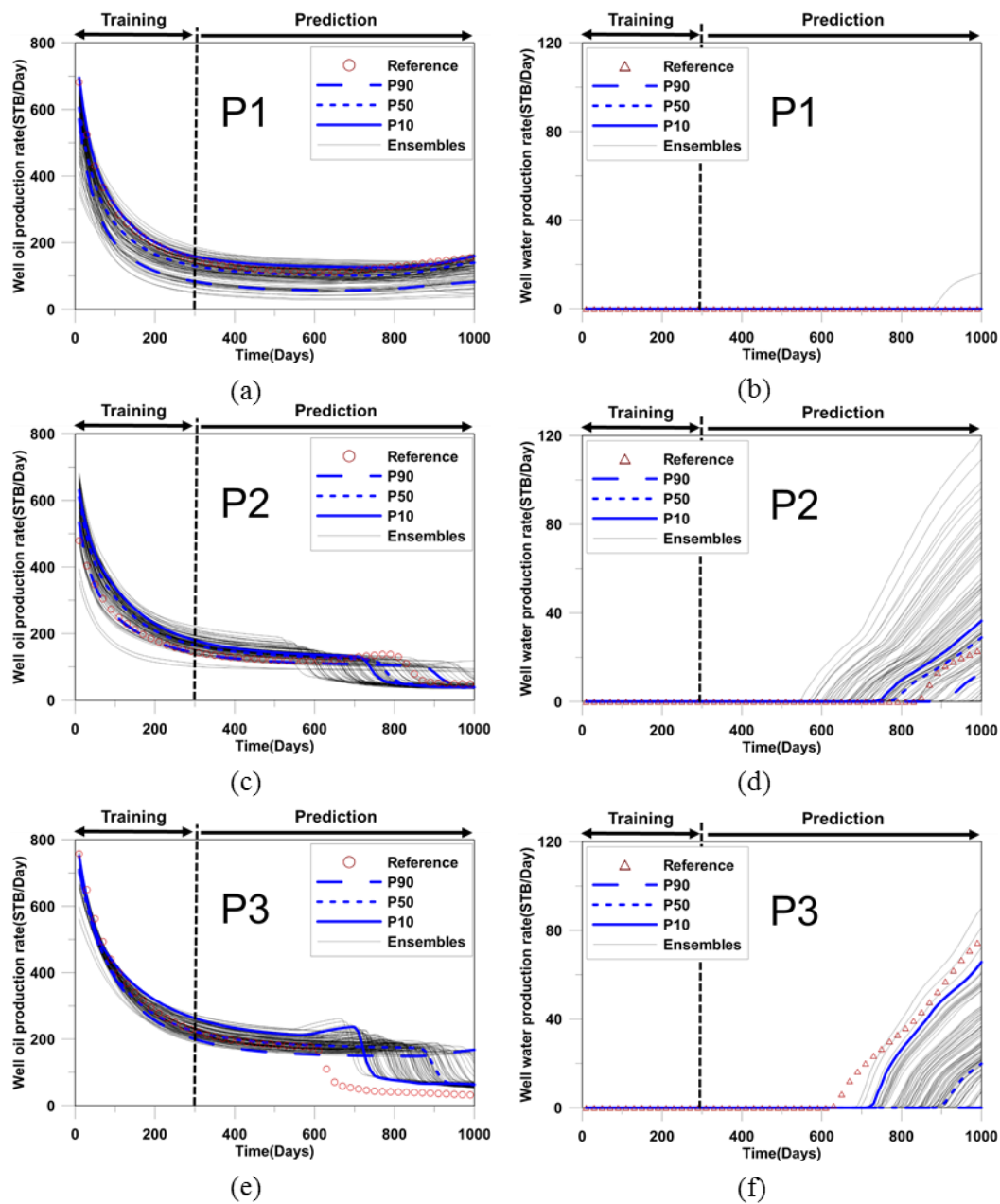
저류층 특성화를 위해 기존의 인공신경망 기법을 적용한 연구에서는 격자 별 유체투과도 값이 아니라 구역별로 평균 유체투과도 값을 균질(homogeneous)하게 가진다는 가정(Ramgulam, 2006)이나, 생산정 위치의 유체투과도 값만을 예측하는 연구(Shahkarami et al., 2014)가 수행되었다. 이 연구에서는 기존 연구와 달리 격자 별 불균질한 유체투과도 값을 예측하였으며, 생산정의 수에 대한 영향을 최소화 하였다.

이 절에서는 인공신경망 기법과 대비하여 개발한 신경망 기법의 적용가능성을 검증해 보았다. 비교 대상은 구역별 평균 유체투과도 값을 가정한 모델이었으며, Figure 4.20은 균질한 유체투과도 모델에 인공신경망 기법을 적용하여 예측한 유체투과도 분포이며, Figure 4.21은 저류층의 유체투과도 분포에 따른 미래생산추이를 예측한 결과이다.



**Figure 4.20 Log permeability distribution: (a) reference (b) predicted ensemble(P50) from a homogenous model.**





**Figure 4.21 Well production rate of predicted ensembles from homogeneous model:**  
 (a) oil rate of P1 (b) water rate of P1 (c) oil rate of P2 (d) water rate of P2 (e) oil rate of P3 (f) water rate of P3.

구역별 균질한 유체투과도 값을 가정한 모델의 미래 생산추이 예측결과는 3번 생산정을 제외한 1번, 2번 생산정에서는 참조모델의 생산추이가 앙상블들의 불확실성 범위에 포함되었으며, 물돌과 발생시점도 비교적 양호하게 예측하였다. 예측성능에서는 우수하다고 볼 수 있지만 균질 모델은 채널연결성은 사전에 알 수 없는 정보임에도 불구하고 구역의 경계를 사전에 알고 있다는 가정하에 사암과 셰일성 사암지역으로 구분되었다. 실제 현장에서는 이러한 정보를 알 수 없기 때문에 현장에서 채널저류층 연결성을 파악하기 위한 의사결정도구로 활용하기에는 한계점이 있다.

하지만 개발한 신경망 기법을 적용한 모델에서는 트레이닝 이미지를 활용하여 채널연결성의 불확실성을 고려하였으며, 격자 별 불균질한 유체투과도 값을 예측하면서 복잡한 지질학적 구조를 신뢰할 수 있게 모사하였다. Figure 4.22는 균질한 유체투과도 값을 가정한 인공신경망 기법, 격자별 비균질한 유체투과도를 가정한 인공신경망 기법, 심층 오토인코더를 적용한 신경망 기법의 일일 오일 생산량 오차를 박스그림으로 나타낸 것이며, Figure 4.23은 세 기법의 펠드 누적 오일 생산량을 비교한 것이다. 일일생산량 예측 성능에서도 개발한 기법이 우수하였으며, 누적 오일 생산량의 불확실성 평가에서 균질한 모델을 사용한 기법의 불확실성이 참조모델의 누적생산량인 1값을 포함하지만 불확실성의 범위가 개발한 기법에 대비하여 넓은 것을 확인할 수 있다.

마지막으로 세 기법의 연산속도 성능을 비교해보았다. Table 4.7은 세 기법의 연산속도를 비교하여 표로 나타낸 것이다. 예측 출력 값이 5개에 불과한 균질모델에서의 연산시간은 267.9초인 반면 비균질한 모델에

인공신경망을 적용할 경우 예측 출력 값이 2,500 개로 늘어남에 따라 연산시간이 12,631초로 급격하게 늘어났으며, 성능 또한 저하되었다. 반면, 개발한 기법은 연산시간이 864.5초로 2,500개의 자료를 100개의 자료로 압축 변환하면서 격자 자료 예측시간을 크게 줄였고, 인공신경망 기법과 대비하여 약 11배 이상의 연산속도를 보여 예측을 위한 시간비용을 감소시켰다.

**Table 4.7 Comparison of average computation time**

<b>Method</b>	<b>Computation time(sec)</b>
Homogeneous ANN	267.9
Heterogeneous ANN	12,631
Proposed NN using autoencoder	864.5

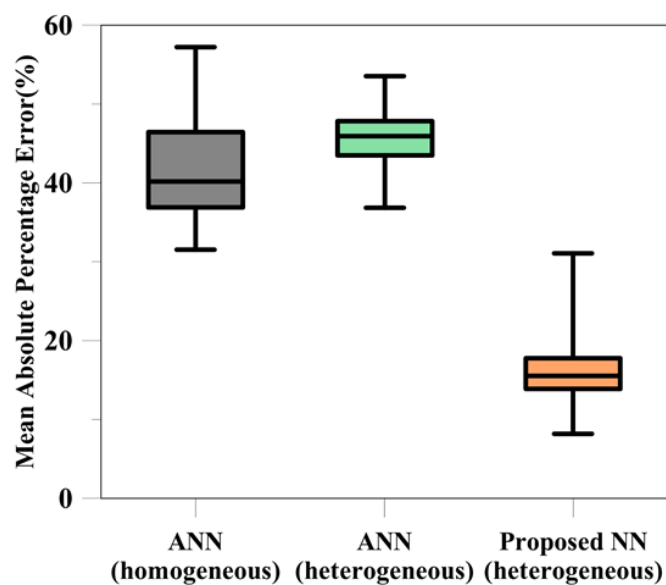


Figure 4.22 Comparison of boxplot of daily well oil productions error.

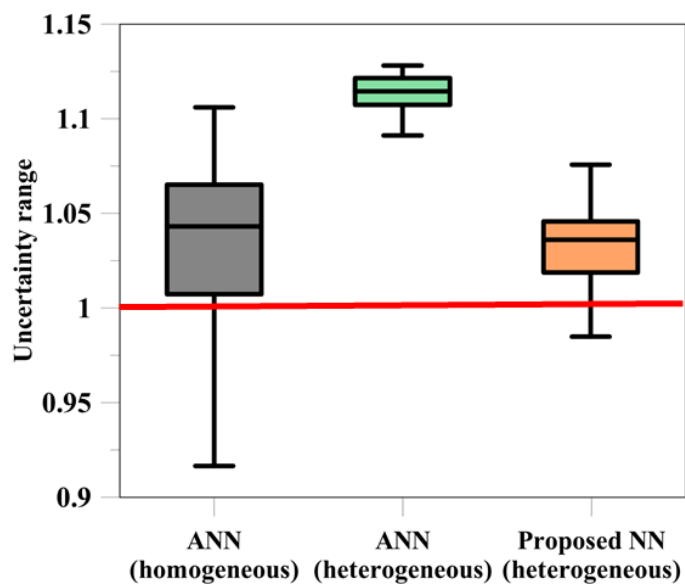


Figure 4.23 Comparison of boxplot of uncertainty range of cumulative oil productions at 1,000 days.

## 5. 결론

이 연구에서는 저류층 특성화를 통한 미래생산추이 예측을 위하여 정적자료인 유체투과도와 동적자료인 생산정 별 오일, 물 생산이력을 통합 분석할 수 있는 심층 오토인코더를 결합한 신경망 기법을 개발하였다. 개발한 신경망 기법의 성능은 참조모델을 포함한 경우와 그렇지 않은 경우의 학습에 대해 일반 신경망기법과 비교하여 생산이력과 유체투과도 분포의 예측정확도 및 연산시간을 평가하였다. 이러한 연구를 통해 도출한 결론은 다음과 같이 요약할 수 있다.

1. 개발한 심층 오토인코더를 적용한 신경망 기법은 기존의 균질한 유체투과도 모델보다 더 복잡한 지질학적 불확실성을 규명할 수 있었다. 다중 출력 예측 문제에서 발생할 수 있는 인공신경망의 연산속도 및 성능 저하를 심층 오토인코더의 자료 압축변환 과정을 통해 보완할 수 있었다. 또한 생산정의 생산이력과 유체투과도의 분포의 비선형적 상관관계를 학습하기 전에 심층 오토인코더를 이용하여 사전에 불균질한 유체투과도의 특성을 학습함으로써 기존의 인공신경망 기법과 달리 신경망을 이용한 비선형관계 학습과정에서도 채널 분포 특성을 보존할 수 있었다.

2. 개발한 신경망 기법은 참조모델의 채널 분포 특성을 보존하면서도 생산정의 미래생산추이 예측에서 우수함을 보였다. 인공신경망 기법은 채널연결성은 모사하였다고 볼 수 있지만 채널저류층의 유체투과도 분포

특징을 유지하지 못하여 참조 모델이 포함된 학습의 경우에도 32%의 생산이력 오차가 발생하였으며, 참조모델이 포함하지 않은 학습의 경우 생산이력의 오차가 45.9%까지 증가하여 미래생산추이 예측에 실패하였다. 반면, 개발한 신경망 기법은 채널연결성을 모사하면서도 유체투과도의 분포 특징을 보존하였기 때문에 참조모델을 포함한 학습의 경우 생산이력 오차가 8.6%였으며, 포함하지 않은 경우에도 오차가 15.6%로 일반 신경망기법에 대비하여 1/3수준으로 매우 양호하였다. 또한 생산정의 물돌과 발생 시점을 양호하게 예측하였다.

3. 개발한 신경망기법은 심층 오토인코더를 이용하여 자료를 압축 변환할 수 있기 때문에 연산속도가 빠르다. 2,500개의 불균질한 유체투과도를 예측하기 위해 인공신경망 기법을 이용할 경우 연산시간이 12,631초가 소요된 반면, 개발한 신경망 기법은 2,500개의 자료를 100개의 자료로 압축하여 학습함으로써 연산시간이 864.5초가 소요되었다. 개발한 신경망 기법은 기존 인공신경망 기법과 대비하여 약 14배 빠른 연산속도를 보여 시간 비용을 크게 절약하였다.

따라서 개발한 심층 오토인코더 기반의 신경망기법은 스케일이 다양한 생산정 별 생산이력과 저류층 물성의 불균질도를 통합하여 신뢰도 높은 미래생산추이 추정이 가능하므로 향후 저류층 개발 및 평가를 위한 합리적인 생산설계 과정에서 의사결정도구로 활용될 수 있을 것이다.

## 참고문헌

이경복. 2014. 거리기반 앙상블스무더를 이용한 채널저류층 특성화. 공학박사학위논문, 서울대학교, 서울.

임서진. 2016. On the Study of a Robust History Matching of Facies Model by Distance-based Method and Ensemble Kalman Filter. 공학석사학위논문, 서울대학교, 서울.

최종근. 2013. *지구통계학*. 서울: (주)시그마프레스.

Araghinejad, S. 2014. *Data-Driven Modeling: Using MATLAB in Water Resources and Environmental Engineering*, Vol 67. Dordrecht, The Netherlands: Springer.

Bengio, Y., Lamblin, P., Popovici, D. and Larochelle, H. 2007. Greedy Layer-Wise Training of Deep Networks. In *Advances in Neural Information Processing Systems*, ed. Schölkopf, N., Platt, J. and Hoffman, T., Chap 19, 153-160. Massachusetts, USA: MIT press.

Centilmen, A., Ertekin, T. and Grader, A. S. 1999. Applications of Neural Networks in Multiwell Field Development. Presented at SPE Annual Technical Conference and Exhibition, Houston, Texas, USA, 3-6 October. SPE-56433-MS. <http://dx.doi.org/10.2118/56433-MS>.

Cullick, A. S., Johnson, W. D., and Shi, G. 2006. Improved and More Rapid History Matching With a Nonlinear Proxy and Global Optimization. Presented at the SPE Annual Technical Conference and Exhibition, San Antonio, Texas, USA, 24-27 September. SPE-101933-MS. <http://dx.doi.org/10.2118/101933-MS>.

Daly, C. 2005. Higher Order Models Using Entropy, Markov Random Fields, and Sequential Simulation. In *Geostatistics Banff 2004*, ed. Leuangthong, O. and Deutsch, C.V., Chap. 22, 215-224. Dordrecht, The Netherlands: Springer.

Doraisamy, H., Ertekin, T. and Grader, A. S. 1998. Key Parameters Controlling the Performance of Neuro-Simulation Applications in Field Development. Presented at SPE Eastern Regional Meeting, Pittsburgh, Pennsylvania, USA, 9-11 November. SPE-51079-MS. <http://dx.doi.org/10.2118/51079-MS>.

Hagan, M. T., Demuth, H. B. and Beale, M. H. 1996. *Neural Network Design*. Boston: PWS publishing company.

Hinton, G., Osindero, S. and Teh, Y. 2006. A Fast Learning Algorithm for Deep Belief Nets. *Neural Computation* **18** (7): 1527-1554. <http://dx.doi.org/10.1162/neco.2006.18.7.1527>.

Jafarpour, B. and Mclaughlin, D. B. 2007. History Matching with an Ensemble Kalman Filter and Discrete Cosine Parameterization. Presented at SPE Annual Technical Conference and Exhibition. Anaheim, California, USA, 11-14 November. SPE-108761-MS. <http://dx.doi.org/10.2118/108761-MS>.

Jafarpour, B. and Mclaughlin, D. B. 2009. Estimating Channelized-Reservoir Permeabilities with the Ensemble Kalman Filter: The Importance of Ensemble Design. *SPE J.* **14**(2): 374-388. SPE-108941-PA. <http://dx.doi.org/10.2118/108941-PA>.

Korjani, M. M., Popa, A. S., Grijalva, E., Cassidy, S. and Ershaghi, I. 2016. Reservoir Characterization Using Fuzzy Kriging and Deep Learning Neural Networks. Presented at SPE Annual Technical Conference and Exhibition, Dubai, UAE, 26-28 September. SPE-181578-MS. <http://dx.doi.org/10.2118/181578-MS>.



Li, D. and Yu, D. 2014. Deep Learning: Methods and Applications. *Foundations and Trends® in Signal Processing* **7** (3–4): 197-387. MSR-TR-2014-21. <http://dx.doi.org/10.1561/20000000039>.

Lechner, J. P. and Zangl, G. 2005. Treating Uncertainties in Reservoir Performance Prediction with Neural Networks. Presented at SPE Europec/EAGE Annual Conference, Madrid, Spain, 13-16 June. SPE-94357-MS. <http://dx.doi.org/10.2118/94357-MS>.

Mitchell, T. M. 1997. *Machine Learning. 1st edition*. New York: McGraw-Hill.

Maltarollo, V. G., Honório, K. M. and Silva, A. B. F. 2013. Applications of Artificial Neural Networks in Chemical Problems. In *Artificial Neural Networks - Architectures and Applications*, ed. Suzuki, K. Chap. 10, 203-223. Rijeka, The Croatia: InTech.

Ramgulam, A. 2006. *Utilization of Artificial Neural Networks in the Optimization of History Matching*. MS thesis. Pennsylvania state University. Pennsylvania. USA.

Reis, L. 2006. Risk Analysis with History Matching Using Experimental Design or Artificial Neural Networks. Presented at SPE Europec/EAGE Annual Conference, Vienna, Austria, 12-15 June. SPE-100255-MS. <http://dx.doi.org/10.2118/100255-MS>.

Sampaio, T. P., Ferreira, V. J. M. and Neto, A. D. S. 2009. An Application of Feed Forward Neural Network as Nonlinear Proxies for Use during the History Matching Phase. Presented at Latin American and Caribbean Petroleum Engineering Conference, Cartagena de Indias, Colombia, 31 May-3 June. SPE-122148-MS. <http://dx.doi.org/10.2118/122148-MS>.

Shahkarami, A., Mohaghegh, S. D., Gholami, V. and Haghghat, S. A. 2014. Artificial Intelligence (AI) Assisted History Matching. Presented at SPE Western North American and Rocky Mountain Joint Meeting, Denver, Colorado, 17-18 April. SPE-169507-MS. <http://dx.doi.org/10.2118/169507-MS>.

Silva, P. C., Maschio, C. and Schiozer, D. J. 2007. Use of Neuro-Simulation Techniques as Proxies to Reservoir Simulator: Application in Production History Matching. *Journal of Petroleum Science and Engineering* **57**(3-4): 273-280. <http://dx.doi.org/10.1016/j.petrol.2006.10.012>.

Tavakoli, R., Srinivasan, S. and Wheeler, M. 2013. Rapid Updating of Stochastic Models Using an Ensemble Filter Approach. Presented at SPE Reservoir Simulation Symposium, The Woodlands, Texas, USA, 18-20 February. SPE-163673-MS. <http://dx.doi.org/10.2118/163673-MS>.

Zupan, J. 1994. Introduction to Artificial Neural Network (ANN) Methods: What They Are and How to Use Them, *Acta Chimica Slovenica* **41**(3): 327-352.

## **Abstract**

# **Characterization of Channelized Reservoir Using the Neural Network Incorporated with Deep Autoencoder**

Seongin Ahn

Department of Energy System Engineering

The Graduate School

Seoul National University

This study reliably predicts the future production of a heterogeneous reservoir by analyzing fluid permeability and production history using neural network technique with deep autoencoder. The artificial neural network - based history matching simplifies the nonlinear relationship of static and dynamic data, so it is difficult to preserve the heterogeneous distribution characteristics of physical properties even if the matching efficiency of production history is high. Due to the failure of this, the consistency of the production trend prediction will not ensure after the history matching. This study proposed a neural network system using a deep autoencoder and learned nonlinear correlation of production history and fluid permeability distribution. The target reservoir is a channel reservoir with heterogeneous fluid permeability with three production wells and one water injection well. The performance of the proposed technique was compared with the

artificial neural network technique for training with and without the target reservoir, and the prediction accuracy of the production history and the fluid permeability distribution and computation time were evaluated. In both cases, the fluid permeability and channel connectivity of the sandstone were closely predicted. However, the artificial neural network technique failed to predict the fluid permeability distribution of the shaly sandstone, and the daily oil production error was 32%. On the other hand, the proposed neural network technique predicted the fluid permeability of the shaly sandstone well, and the daily oil production error decreased to 8.6%. Also, even if the target reservoir is not included, it has 15.6% error which means that error is 1/3 level even though the computation time is about 14 times faster than the artificial neural network technique. In the heterogeneous distribution prediction, the prediction results of the training of the proposed neural network were similar to the target reservoir in both permeability and channel distribution although the artificial neural network failed to predict the distribution of permeability of sandstone channel. The neural network technique based on the deep autoencoder can be used as a useful tool for decision making of the production design process in reservoir engineering because it can estimate the production trend with high reliability by integrating production history with various scales and heterogeneity of reservoir properties.

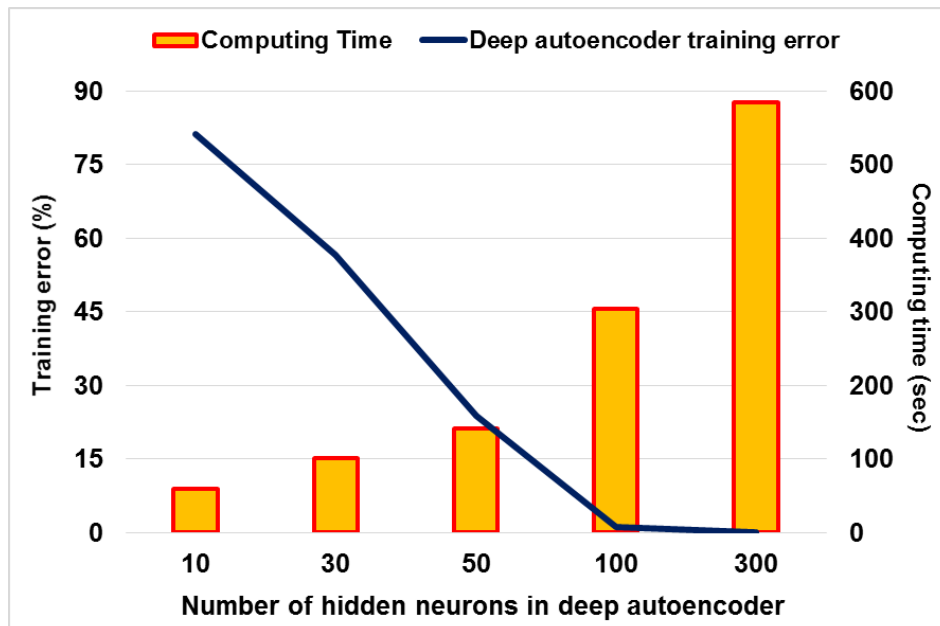
**Keywords:** neural network, artificial neural network, deep autoencoder, channel reservoir

**Student Number:** 2015-21320

## 부록 A. 심층 오토인코더의 은닉층 뉴런수의 결정

심층 오토인코더는 자료를 사용 목적에 따라서 압축하거나 확장할 수 있다. 많은 양의 자료를 처리하면서 연산 시간이 오래 걸리고 성능이 저하될 때 자료를 압축 변환하게 되면 연산속도를 증가시켜 시간비용을 크게 저감시킬 수 있다. 하지만 압축 변환이 가능하다고 해서 원본 데이터의 성질을 잃어버릴 만큼 압축시킬 경우 문제가 될 수 있다. 이에 본 연구에서는 오토인코더의 은닉층 뉴런 수에 따른 연산속도와 성능을 검증해 보았다.

Figure A.1은 오토인코더의 은닉층 뉴런 수에 따른 검증 결과이다. 뉴런의 수가 줄어들수록 연산시간이 감소하는 것을 확인할 수 있다. 하지만 뉴런의 수가 100개 이하로 줄어들 경우 원본자료의 특징을 잃어버리면서 재구현하는 성능이 급격하게 감소하는 것을 확인할 수 있다. 뉴런의 수가 많을수록 자료의 재구현 성능은 높아지지만 100개인 경우 300개의 경우와 성능의 차이는 거의 나지 않지만 연산시간은 2배 이상 차이가 났다. 이에 따라 본 연구에서는 은닉층의 뉴런 수를 100개로 설정함으로써 연산속도도 빠르면서 성능 면에서 발생한 수 있는 원본 자료 재구현에 따른 오류의 불확실성을 크게 감소시켰다.



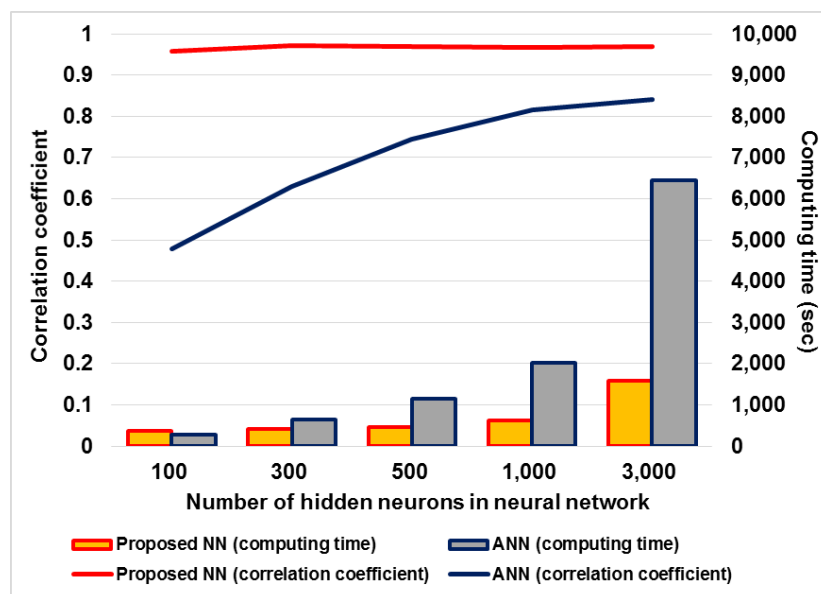
**Figure A.1 Error variance and computing time depending on the number of hidden neurons in deep autoencoder.**

## 부록 B. 신경망 기법의 은닉층 뉴런수에 따른 성능

신경망 기법을 이용하여 입력층과 출력층의 비선형관계를 예측하기 위해서 은닉층의 뉴런 수를 적절하게 설정하는 것은 중요한 과정이다. 일반적으로 은닉층의 뉴런 수는 학습의 안전성을 위해 입력층과 출력층의 뉴런수 보다 많아야 하지만 뉴런의 수가 증가함에 따라 연산시간이 증가하게 된다. 이를 바탕으로 이 연구에서는 생산이력과 유체투과도 분포의 비선형관계 예측을 위한 신경망 구조 구축과정에서 은닉층의 적절한 뉴런 수를 설정하기 위하여 뉴런 수에 따른 성능을 평가해 보았다. 인공신경망 기법과 개발한 심층 오토인코더를 결합한 신경망 기법의 연산시간과 예측성능을 평가하였으며, 예측성능은 참조모델과 예측한 모델의 유체투과도 상관관계(Correlation)를 기준으로 평가하였다. 평가를 위하여 두 신경망 기법의 학습 수(epoch)를 포함하여 은닉층의 뉴런 수를 제외한 나머지 인자들은 동일하게 설정하였으며, 10개의 예측모델에 대한 연산시간과 상관계수를 비교하였다.

Figure B.1은 신경망 기법의 은닉층 뉴런 수에 따른 성능 비교 결과이다. 인공신경망 기법의 경우 출력층의 뉴런 수가 2,500개이므로 은닉층의 뉴런수가 500개 미만일 경우 상관계수가 0.75미만으로 예측성능이 좋지 못하였다. 뉴런수가 1,000개일 경우 상관계수가 0.81이었으며, 3,000개일 경우 0.84까지 성능이 향상 되었다. 하지만 연산시간이 1,000개의 경우 약 2,000초, 3,000개의 경우 약 6,500초까지 증가하는 추세를 보였다. 반면, 개발한 신경망 기법은 입력층이 180개, 출력층이 100개의 뉴런을 가지고 있기 때문에 은닉층의 뉴런 수가 100개일 경우 0.94의 상관계수 값을 보였으며, 나머지 경우에는 상관계수가

0.97로 매우 양호한 성능을 보였다. 또한 뉴런수가 증가하더라도 연산시간만 증가할 뿐 예측성능이 향상되지 않았다. 연산시간 측면에서도 2,500개의 유체투과자료를 100개의 자료로 압축 변환하는 학습과정에서 321.63초의 시간이 소요되었으며, 비선형관계 예측을 위한 신경망학습 연산시간의 증가는 인공신경망과 대비하여 작은 것을 확인하였다. 이러한 결과를 바탕으로 개발한 신경망 기법이 인공신경망과 대비하여 성능측면에서도 우수할 뿐만 아니라 연산속도도 빠른 것을 확인하였다. 그리고 신경망 구조를 위한 은닉층의 적절한 뉴런 수는 인공신경망의 경우 사용자의 시간비용 고려에 따라 1,000~3,000개 사이를 설정하는 것이 적절하며, 개발한 신경망 기법의 경우 300개 이상일 경우는 성능의 향상 없이 연산시간이 증가하였기 때문에 300개로 설정해 주는 것이 적절하다.



**Figure B.1 Comparison of correlation coefficient and computing time between ANN and proposed NN.**